Developing Solution-focused Sustainability Assessments

Michiel Christiaan Zijp

Author: Michiel C. Zijp
Cover: "Speelruimte" door Wout Zijp, www.woutzijp.nl
Lay-out: DTProost - DTP Studio
Printing: Ecodrukkers Nieuwkoop
This thesis is printed fully CO2 neutral on FSC certified biotop paper with the use of eco-ink and a minimum of hazardous substances.
ISBN: 9789492477521

Developing Solution-focused Sustainability Assessments. PhD-thesis, Radboud University, Nijmegen.

Developing Solution-focused Sustainability Assessments

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 3 juli 2017, om 10:30 uur precies

door

Michiel Christiaan Zijp

geboren op 28 december 1980
te Zeist

Promotoren
Prof. dr. L. Posthuma
Prof. dr. M.A.J. Huijbregts

Copromotor
Dr. R. Heijungs (Universiteit Leiden, Vrije Universiteit Amsterdam)

Manuscriptcommissie
Prof. dr. ir. A.J. Hendriks
Prof. dr. A.M. Breure
Prof. dr. A. Tukker (Universiteit Leiden)

Paranimfen
Dr. A. Hollander
Dr. A. Vonk Noordegraaf

Voor mijn moeder

Table of contents

1 Introduction

1.1 Sustainable development

In 2016 the Sustainable Development Goals (SDGs) came into force (UN, 2015a). With this "2030 agenda for sustainable development", the 194 member states of the United Nations acknowledged the vast challenges humanity faces. Its 17 goals, subdivided in 169 targets, cover a broad range of issues, varying from human rights – such as ending poverty and hunger, and increasing gender equality – to environmental issues – such as reducing climate change and biodiversity loss – and sustainable economic growth. Some authors criticized the SDGs for being too broad, hampering action for change (Paulson, 2015). Others stress the inclusiveness of the goals, highlighting that the challenges are interconnected, with possible trade-offs, and should be approached as such (Hopwood et al., 2005; Kates et al., 2001). For example, a program to reduce poverty should not enhance climate change and vice versa (Renton, 2009).

Raworth (2012) presented a conceptual model that expresses the connection between social and environmental issues (Fig. 1.1) (Dearing et al., 2014). The model summarizes the view that our human rights must be met (social foundation, the inner circle in Fig. 1.1) within the carrying capacity of our environment (environmental ceiling, the outer circle in Fig. 1.1). The space between the circles is called "the safe and just space for humanity" by these authors. Efforts to reach this space, and stay there, are defined sustainable

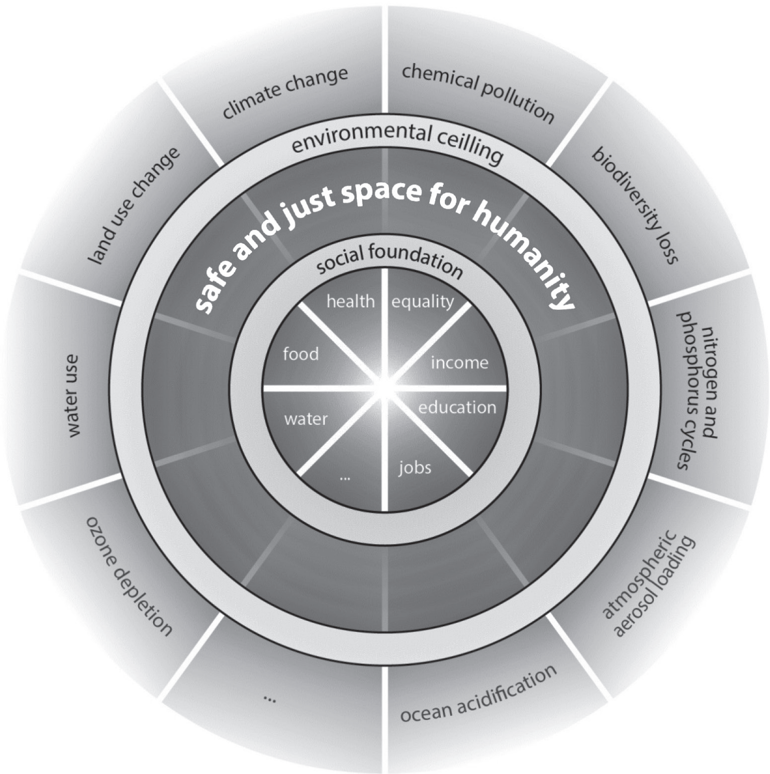


Fig. 1.1 A safe and just space for humanity (adjusted from: Dearing et al. (2014); Raworth (2012)) is the situation in which basic human needs are met within the carrying capacity of the environment.
Aiming for this situation requires consideration of social and environmental issues of which some examples are shown in the figure.

development (ibid.). The SDGs show that both the social foundation and the environmental ceiling encompass many different issues of which an exemplary selection is presented in Fig. 1.1. This means that an assessment on breaching or reaching the safe and just space for humanity requires multiple metrics from both the environmental as the socio-economic realms.

Defining sustainable development. The definition of sustainable development can be discussed on the level of general description and principles and on the level of operationalization. Sustainable development is well known for its many interpretations. It's definition is discussed ever since it gained attention via the World Conservation Strategy (IUCN et al., 1980) and later the Brundtland report (Brundtland, 1987), which defined sustainable development as 'meeting the needs of the present without compromising the ability of future generations to meet their needs'. Especially the inclusion of economic growth is subject of debate. Proponents believe economic growth is the solution to reduction of poverty of all (Dollar and Kraay, 2000), while others argue that economic growth is the main cause of environmental degradation and does not solve poverty (Daly, 1993). This discussion, started in the years after the Brundtland report (Daly, 1993), is still ongoing. For example as critical note on the SDGs (Hickel, 2015) which explicitly includes reference to quantitative economic growth: "Promote sustained, inclusive and sustainable economic growth, …" (Goal 8, p.19) and "…at least 7 per cent gross domestic product growth per annum in the least developed countries" (Target 8.1, p.19). The definition of sustainability and sustainable development are as broad as conceptualized in Fig. 1.1: (directing towards) the situation in which social needs are met and environmental ceilings are not crossed. As a further specification: sustainability is a situation, while sustainable development is a direction. Efforts to reach a sustainable situation contribute to sustainable development. Those efforts can take place on all organizational levels: on the global level stimulated by choices, decisions and programs of international governmental organizations and multi-national companies, up till the local level, e.g. triggered by life style choices of individuals. Moving to- or staying within the safe and just space requires action of every organizational level everywhere (Hajer et al., 2015).

Next to its definition, different authors proposed different sets of principles to specify sustainable development (e.g. Daly (1990); Haughton (1999); Pintér et al. (2012); Robèrt et al. (2002)). Furthermore, different views exist on interchangeability of social, environmental and economic values in decision-making (strong versus weak sustainability (Neumayer, 1999; Özdemir et al., 2011)), while also, different views exist on the required efforts to establish a transition towards sustainable development – ranging from implementing readily available measures, to considerations on structural societal transformation (Hopwood et al., 2005). This proliferation of interpretations and views on sustainable development can lead to pessimism on the usefulness and meaningfulness of the term. On the other hand, the use of the term sustainable development in its wide variety of possible interpretations might also serve a goal (Laniak et al., 2013). It appears to help in the initial aligning of the viewpoints of different people with different views, which is a starting point for cooperation between otherwise separated parties (Velten et al., 2015). However, operationalizing sustainable development in the context of decision-making after getting different parties together requires specific choices, and it is here that sustainable development occurs, and can or cannot become materialized and thereupon also measurable.

Translating sustainable development in principles or goals is still relative 'safe', because those can be very comprehensive, thereby including everybody's views. Assessment of sustainable development, as needed to support decision-making in practice, requires however very specific translations of the sustainable development principles and goals: i.e. indicators and metrics through which the status (sustainability) or the process (sustainable development) can be characterized. Such a characterization may thereby involve metrics from the environmental or social sciences (outer and inner circle), but also value judgments to define an 'undesired' status for any of those metrics. Because the inclusion of all interconnected sustainability issues in an assessment is both unfeasible and unpractical, choices need to be made on which issues are critical in defining the undesired status (and thus be included), which issues are of key interest from the viewpoint of stakeholders, and whether trade-offs may be present. Because of the multi-dimensional nature of the sustainable-development concept and the element of 'value judgement', the sustainability assessment methods used to analyze a situation should be selected tailored not only to the physical situation, but also to the views on sustainable development of the people involved. In practice however, the selection of methods for sustainability assessment appears to be largely driven by the availability of expertise and data instead of a conscious choice related to the situation (Gasparatos and Scolobig, 2012). In order to enhance the utility of sustainability assessments for sustainable-development oriented decision-making, the process of method selection has to become more explicit, inclusive and transparent.

Sustainable development and wicked problems. Although the model of Raworth (Fig. 1.1) presents the sustainable development challenge at a planetary level, it is a challenge that includes all spatial scales. That is, ecosystems are under the influence of human induced stress at all scales (Maxwell et al., 2016). Some authors even speak of a nearly global 'anthropogenic biosphere' (Ellis, 2015) as characterization of today's environmental status. Most environmental issues take place on a regional rather than planetary scale, because they are a product of local combinations of stressors and local environmental ceilings. The latter depends on the carrying capacity and the regeneration capability of the local ecosystems (Dearing et al., 2014; Teah et al., 2016). Local or regional ceilings, when transgressed in different regions, can thereupon act as a cumulative driver of global change (Bernhardt et al., 2017; Rockström et al., 2009a; Rockström et al., 2009b).
Assessing or steering towards sustainable development can be complex for more reasons than the already mentioned interacting environmental and social issues and the different spatial scales. Choices for sustainable development often involve conflicting goals (e.g. a growing economy and reducing climate change) and multiple perspectives of stakeholders on what sustainable development is and how to achieve it. As a consequence, there is often no single optimum to strive for and developments that are valued sustainable by one group of people can be valued disastrous by others. For example, health councils can make a plea for increased consumption of fatty fish because of positive health effects of fish fatty acids, while animal-care organizations plea to reduce fish catch because of animal wellbeing arguments (Hollander et al., in prep). These are typical features of so-called 'wicked problems', as defined by Rittel and Webber (1973). Finding common agreed solutions for wicked problems requires iterative explorations of solution scenarios, including the engagement of stakeholders (Roberts, 2000). Various authors envision an important and changing role for scientists in the societal process of dealing with wicked problems and proposed sustainability science as a new discipline to explore this challenge (Clark and Dickson, 2003; Kates et al., 2001; Sala et al., 2012b). Sustainability science is considered a relative new discipline that "seeks to understand interactions between nature and society" (Kates et al., 2001) and that contributes to solutions for sustainable development in societal complex processes (Clark and Dickson, 2003; Sala et al., 2012b). It is then envisioned that scientists are part of the self-learning process towards solutions for these wicked problems, in interaction with the stakeholders of the respective problem (De Ridder et al., 2007; Pooley et al., 2013; Sala et al., 2013; Sala et al., 2012b; Thabrew et al., 2009).

The orientation of sustainability science towards solutions is a specific feature of this discipline. The shift towards more solution-focused approaches to support decision-making has been made in different policy context, and is today supported by major jurisdictions, advised by scientific advisory boards:

- Commissioned by the U.S. Environmental Protection Agency, the National Academy of Sciences of the United States, (U.S. NAS, 2009) proposed "solution-focused risk assessment", which implies exploring risk reduction scenarios before, rather than after, an assessment of risks; the motive to suggest the solution-focused approach is the conclusion that there is a need for a higher utility of risk assessments.

- The European Commission (EC) hints at using a solution-focused approach for sustainability optimization via 'nature-based solutions' (EC, 2015) and finances a major European research project that is based on the solution-focused approach, addressing solutions to chemical pollution of water resources (Brack et al., 2015).
- the United Nations launched the Sustainable Development Solutions Network to invite research institutes and universities to contribute in finding solutions for reaching the Sustainability Development Goals (UN, 2015b).

However, currently, sustainability assessments often still tend to focus on the identification of the type and magnitude of (risk) problems instead of exploring possible solutions for sustainable development (Clark and Dickson, 2003; Waas et al., 2014).

In summary, today's challenge for sustainability assessments that contribute to decision-making towards reaching the SDGs is that they must be tailored to the context, cover multiple differing social and environmental issues and be solution focused.

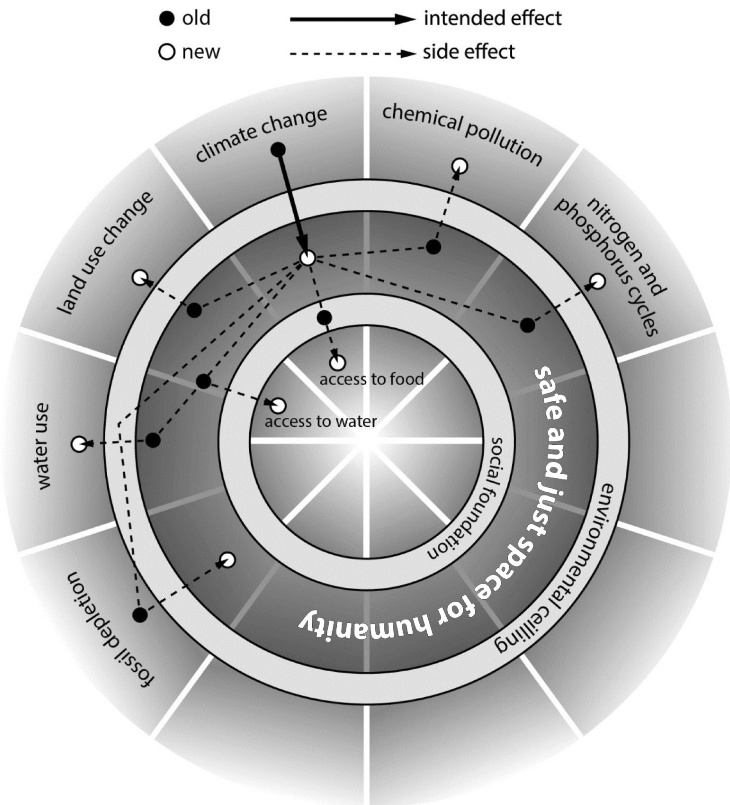


Fig. 1.2 Hypothetical illustration of the multidimensional character of sustainability questions: replacing fossil fuels with biofuels as solution to net reduction greenhouse gas emissions, comes with the use of land, water, pesticides and might also conflict with food security. These are thus all interconnected issues that require consideration before choosing for this solution or an alternative solution scenario. Black dots: original situation; white dots: future situation. Black arrow: primary trajectory, solving a problem. Dotted arrow: associated trajectories.

1.2 Sustainability assessments

Sustainability assessment methods. Sustainability assessments are efforts to support sustainable development with insights in the consequences of future or past human activities from a multi-disciplinary point of view (Bond et al., 2012; Waas et al., 2014). The demand for support on decisions towards sustainable development has resulted in a plethora of methods that claim to provide this support. Every method, such as the Dow Jones Sustainability Index, the Product Environmental Footprint and the Eco-cost Value Ratio, offers a specific translation and operationalization of the concept of sustainable development into pertinent metrics. These translations differ, because, firstly, like mentioned above, people tend to have different views on sustainable development (Velten et al., 2015; Winnebeck, 2011) and secondly, different situations, e.g. spatial scales, require different insights in sustainable development (De Ridder et al., 2007; Wrisberg et al., 2000). It is not feasible and unpractical to include all possible sustainability issues in every assessment (Steinmann et al., 2016). Therefore, the current sustainability assessment methods and results include a selection of issues. Hence, a choice for a sustainability assessment method comes with choices that determine where to focus time and effort on, and thus how sustainable development is specified, - e.g., which themes (such as climate change, child labor, economic benefits) should be included in the assessment. As a consequence, in order for a sustainability assessment to be supportive for decision-making towards a more sustainable society, selected methods should be chosen such that they link to the problem and its context (Harder, 2015). When stakeholders and decision makers are not included in the process of method selection there is a chance they do not recognize their view on the problem and their view on sustainable development in the results. Hence, the assessment results do not lead to a decision, but to a discussion on what is important in terms of sustainability. Discussion on the definition of sustainability before the assessment takes place and participation of the stakeholders in the process of method selection could increase the chance of results that are supported and used in the decision-making process (Lind et al., 1990; Lind and Tyler, 1988). In practice however method selection is largely driven by experience with method application and data availability and rarely based on a transparent definition of sustainable development (Gasparatos and Scolobig, 2012; Little et al., 2016; Sala et al., 2013).

Environmental boundaries. Although many sustainability assessment methods are available to assess sustainable development, ideas on how to assess whether sustainability is reached are scarce. With other words, most methods assess the direction of the development by comparing the impacts of alternative solution scenarios. They do not assess if solutions result in a situation between the social foundation and the environmental ceiling (Fig. 1.1). This may be related to complexities in defining the ceiling and also to the fact that many activities only relate to a small fraction of the set of problems that cause transgressing a ceiling. There is active debate on how to substantiate the conceptually attractive thought of environmental ceilings. So-called environmental boundaries have been proposed to indicate the position of environmental ceilings. Setting environmental boundaries is complex, because it depends on the type of impact with its specific stress-response curves (Dearing et al., 2014), the spatial scale (Brook et al., 2013; Rockström et al., 2009b), uncertainty, interactions with other impacts and risk perceptions (De Schryver et al., 2011). Often, a true environmental boundary (also called a natural threshold), cannot exactly be determined, due to the type of stress-response curve (Fig. 1.3), uncertainty and complex interactions between the different stressors and ecosystems. Therefore, environmental boundaries are derived by merging scientific evidence and value judgements on which risks are acceptable and which are not.

In 1982 Siebert (1982) proposed to define an “eco-space” method to compare the combined human induced impacts at a preset spatial scale with the carrying capacity of the environment at that scale. With the introduction of the concept of planetary boundaries in 2009, this work has again the full attention of scientists and politicians (Rockström et al., 2009a). Next to objective descriptions of biophysical processes, these boundaries are partly normative. They are a product of risk perceptions, i.e. based on the precautionary principle (staying on the safe side) and ambitions, i.e. to protect the stable situation of the Holocene (Rockström et al., 2009a). Depending on the type of stress-response curve, an empiric natural threshold value can be derived for various stress-response relations. A hypothetical example can be found in Fig. 1.3 (solid line, T1). Subsequently, a boundary can be set at the safe side of such a threshold (B1 in Fig. 1.3). In many cases however there is no (clear) natural threshold that can serve as anchor point for the boundary (dashed lines in Fig. 1.3). The environmental boundary is then based on an interpretation of the stress-response curve (shape and position) and normative choices, again to derive a safe boundary. This boundary can subsequently serve as reference point for decision-making. Likewise, boundaries for the social foundation can be based on widely agreed social norms, like described in the millennium goals and the SDGs (Raworth, 2012; UN, 2015a).

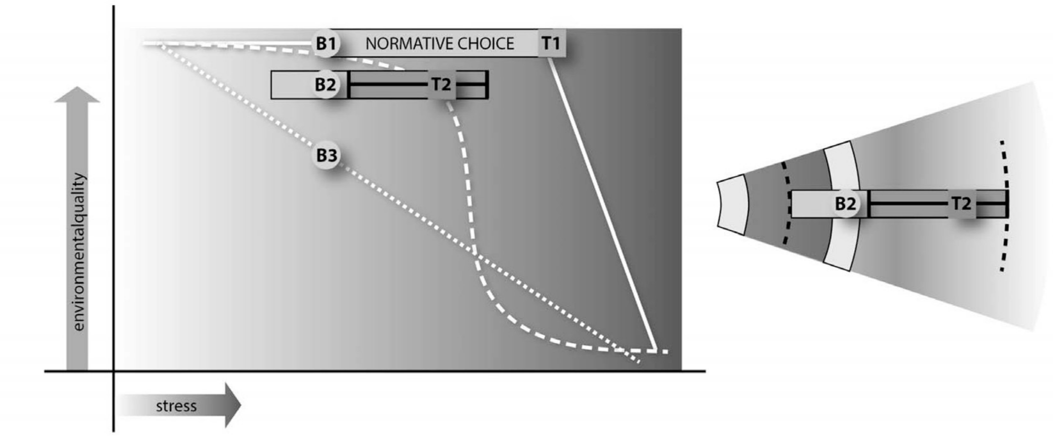


Fig. 1.3 Environmental boundaries (B) used in decision support are set based on stress-response curves, possible natural thresholds (T1) or selected threshold levels (T2) within these curves and normative choices. The dotted lines in the right plot indicate the range in which the boundary can be set based on the sigmoidal response curve.

Subsequently, even when boundaries are derived, their application as object in a decision support process is challenging for two reasons. First, meeting the boundaries depends on many activities (on the planetary scale: all activities), while sustainability assessments regularly focus on just one or a few of them. So, which part of the safe space can be allotted to the different activities, and according to which governance principles (Clift et al., 2017; Nykvist et al., 2013)? Secondly, although human activities result in pressures all over the world and stressors like chemical pollution and eutrophication fit in the definition of drivers of global change (Bernhardt et al., 2017), the responses of social systems and eco-systems are often location-specific. For example, the minimum income to survive (part of the social foundation) is country specific, depending on living costs within a country. Likewise, human impacts on biodiversity are location-specific, depending on the living conditions of the species at a location. As a consequence, global boundaries for many environmental stressors, such as chemical pollution and nutrients loads, need be addressed by aggregating local ones, whereby the set of local boundaries in fact form a distribution of boundaries rather than a single number.

1.3 Goal and research questions

The above sketched situation is that there are needs to consider sustainable development at all spatial scales, translated amongst others in the SDGs, and that the goal of sustainability assessments is to support decision-making that leads to a sustainable situation. These are typical decision procedures that cover multiple social and environmental issues, trade-offs between those issues and conflicting stakeholder perspectives on what is important to be considered and what is not. The intended decision support utility of sustainability assessments in such situations is hampered by:

1. the tendency to focus assessments on problems instead of possible solutions (Clark and Dickson, 2003; U.S. NAS, 2009);
2. method selection that is poorly or not fitting the decision context and the views of the involved stakeholders on sustainable development (Gasparatos and Scolobig, 2012; Waas et al., 2014); and
3. the struggle with boundaries: how do we derive boundaries, especially for those with apparently distributed stressor and vulnerability levels, and how do we use them in governance to forward sustainable development (Pope et al., 2004; Sala et al., 2013)?

The goal of this thesis was to conceptually design, operationalize and test a solution-focused sustainability assessment framework to be used in addressing and solving wicked environmental problems. The operationalization consisted of a procedural and a methodological part and both are illustrated with case studies.

The procedural part focusses on designing and testing a comprehensive approach for executing a sustainability assessment process with a solution-focused approach and a transparent method selection. Within the procedural focus the goal was detailed as follows:

1. *to design and illustrate a solution-focused decision procedure for wicked environmental problems with a clear role for sustainability assessments.*
2. *to design and operationalise a procedure with which a well-informed choice for a sustainability assessment can be made that fits the decision context and can be embedded in the decision procedure as described under 1.*

The methodological part focusses on the utilization of the approaches developed under 1 and 2 to specific situations with man-made stress to ecosystems. The situations concern the quantification of the stress and of the vulnerability of ecosystems to stress, which both vary in space and time. Special attention is paid to chemical pollution as "ignored driver of global change" (Bernhardt et al., 2017), and efforts were made to align chemical pollution with the other stressors considered. Within the methodological focus the sub-goals were detailed as follows:

3. *to quantify the impact of multiple stressors on ecosystems and confront the outcomes of that with different types of regional boundaries.*
4. *to derive and apply vulnerability distributions of ecosystems that can be used to quantify environmental impacts within a region based on both variation in stress as in ecosystem vulnerability.*

1.4 Outline

All chapters are centered around recognized environmental problems, for which the process aspects of sustainability assessments and the scientific challenges (on method selection and stressor and vulnerability quantification) are combined, and tested in case studies. Chapter 2 holds the proposal and illustration of a novel framework for solution-focused sustainability assessment, (SfSA). The framework places sustainability assessments in the context of the other process activities of the SA that have to be performed in order to deal with wicked environmental problems.

Then, chapter 3, shows a literature review that was performed on what choices should be made for a well-informed selection of sustainability assessment method, triggered in the SfSA framework. Here, I propose to develop a sustainability assessment identification key (SAIK) that bridges the gap between the analysis of the problem context and the associated sustainability questions (demand)- and the required methods (supply) sides of sustainability assessments.

In chapter 4, the identification key proposed in chapter 3 is operationalized and applied in a case study on strategic choices between different option to recover resources from domestic waste water in the Netherlands. Also the first four steps of the solution-focused sustainability assessment framework (chapter 2) are applied in this case study.

Chapter 5 presents a novel method that allows quantification of the net toxic pressure caused by multiple human activities within the boundaries of a region. The method is applied on the use of organic substances and pesticides in parts of Europe and explores the use of both policy and natural boundaries for judging net toxic pressures. The novel method results in a chemical footprint, and is thus a chapter oriented on both the stressor-side and the boundary-definition side of the judgements that are needed in a SA.

Chapter 6 presents a novel method in which the variation of vulnerabilities of ecosystems, related to the issue of regional boundaries, is operationally defined. It describes what ecosystem vulnerability distributions (EVDs) are, how they can be derived and how they can be used in stressor identification and ranking. The method was developed, and is illustrated, with a case study on fish assemblages in the state Ohio (U.S.A.).

The findings of this thesis are then brought together in the Synthesis (Chapter 7) and discussed in light of the overall goal: to conceptually design, operationalize and test a solution-focused sustainability assessment framework to wicked environmental problems.

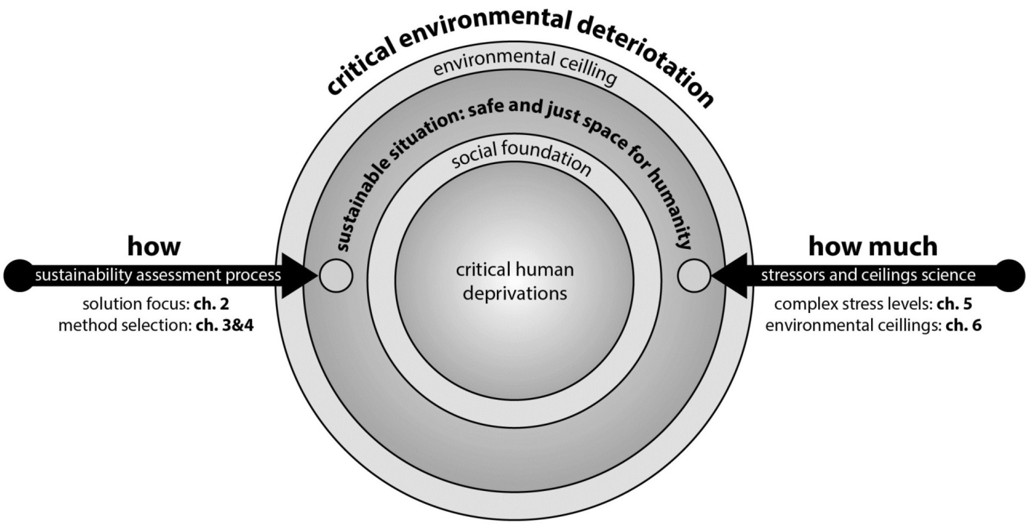
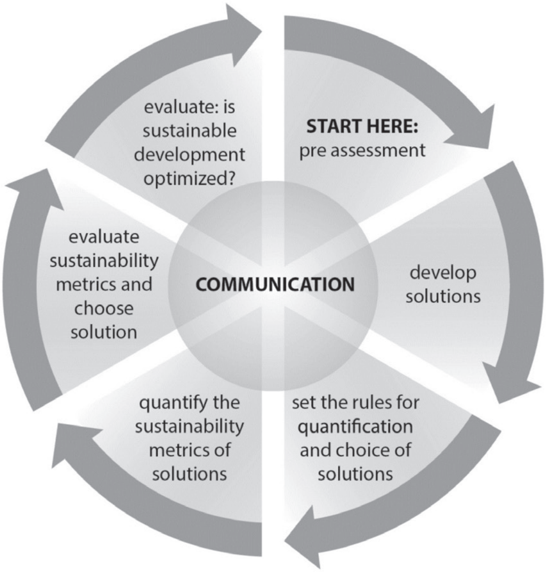


Fig. 1.4 Contribution of this thesis to sustainability assessments that support decision-making that aims for sustainable development.

2 Definition and use of Solution-focused Sustainability Assessment

Authors: Michiel C. Zijp, Leo Posthuma, Arjen Wintersen, Jeroen Devilee, Frank A. Swartjes

Published in: Environment International (2016) 91: 319-331



Abstract

This paper introduces Solution-focused Sustainability Assessment (SfSA), provides practical guidance formatted as a versatile process framework, and illustrates its utility for solving a wicked environmental management problem. Society faces complex and increasingly wicked environmental problems for which sustainable solutions are sought. Wicked problems are multi-faceted, and deriving of a management solution requires an approach that is participative, iterative, innovative, and transparent in its definition of sustainability and translation to sustainability metrics. We suggest to add the use of a solution-focused approach. The SfSA framework is collated from elements from risk assessment, risk governance, adaptive management and sustainability assessment frameworks, expanded with the 'solution-focused' paradigm as recently proposed in the context of risk assessment. The main innovation of this approach is the broad exploration of solutions upfront in assessment projects. The case study concerns the sustainable management of slightly contaminated sediments continuously formed in ditches in rural, agricultural areas. This problem is wicked, as disposal of contaminated sediment on adjacent land is potentially hazardous to humans, ecosystems and agricultural products. Non-removal would however reduce drainage capacity followed by increased risks of flooding, while contaminated sediment removal followed by offsite treatment implies high budget costs and soil subsidence. Application of the steps in the SfSA-framework served in solving this problem. Important elements were early exploration of a wide 'solution-space', stakeholder involvement from the onset of the assessment, clear agreements on the risk and sustainability metrics of the problem and on the interpretation and decision procedures, and adaptive management. Application of the key elements of the SfSA approach eventually resulted in adoption of a novel sediment management policy. The stakeholder participation and the intensive communication throughout the project resulted in broad support for both the scientific approaches and results, as well as for policy implementation.

2.1 Introduction

Risk assessment can be defined as scientific support for decision-making under uncertainty (Yoe, 2011), with risk reduction as ultimate management goal. Recently, the novel concept of Solution-focused Risk assessment (SfRA) was introduced, to improve the utility of risk assessments (Abt et al., 2010; Finkel, 2011; U.S. NAS, 2009). The classical approach (U.S. NAS, 1983) aims to provide insights in current risks and its uncertainty, and suggests refined risk assessment loops to reduce uncertainties, until the results are considered sufficient for transfer to the risk management phase. The solution-focused approach however, explores risk reduction scenarios before, rather than after, the risk assessment. It yields comparative risk levels of the current situation and alternative solution scenarios in a single assessment round. The comparative risk assessment of risk reduction scenarios is followed by selecting the most promising solution scenario and adaptive management loops when needed.

Major jurisdictions support the development and use of solution-focused approaches: (1) Commissioned by the U.S.- Environmental Protection Agency, the National Academy of Sciences of the United States, (U.S. NAS, 2009) proposed the approach. (2) The European Commission (EC) hints at using a solution-focus for sustainability optimization via 'nature-based solutions' (EC, 2015) and finances a major European research project that is based on the solution-focused approach, addressing solutions to chemical pollution of water resources (Brack et al., 2015). And (3) the United Nations launched the Sustainability Development Solutions Network to invite research institutes and universities to contribute in finding solutions for reaching the Sustainability Development Goals (UN, 2015b).

Expanding the use of the 'solution-focused' idea from the realm of risk assessment into that of sustainability assessment logically yields the concept of Solution-focused Sustainability Assessment (SfSA), which we introduce and illustrate in this paper. Risk and sustainability assessments have similarities (Sexton and Linder, 2014; U.S. NAS, 2011), but there are differences. Sustainability assessments usually address multiple metrics, covering the classical sustainability domains of 'people', 'planet' and 'profit/prosperity'. Sustainability science has been proposed as a new discipline in 2001 (Kates et al., 2001). It is considered a solution-focused discipline (Clark and Dickson, 2003; Sala et al., 2012b), geared to support decision-making. Sustainability science further tends to address wicked problems (Stahl and Cimorelli, 2013), for which the main characteristics and coping strategies are summarized in Table 2.1. Contributing to solutions for these problems requires the involvement of multiple disciplines and stakeholders (Pooley et al., 2013; Sala et al., 2012b; Thabrew et al., 2009) and a clear role for scientists in the decision pathways (De Ridder et al., 2007; Marshall et al., 2017; Sala et al., 2013; Spruijt et al., 2014; Thabrew et al., 2009). Just like risk assessments, however, sustainability assessments in practice rather tend to focus on the identification of the type and magnitude of a (risk) problem than on exploring low-risk or high-sustainability solutions (Clark and Dickson, 2003; Waas et al., 2014).

Sustainability-oriented management solutions are especially needed when a ban on an activity is based on the evaluation of a single risk metric, whilst the same activity would be allowed or even stimulated from a multi-metric sustainability assessment. As one example, Riding et al. (2015) analysed scientific, regulatory and socioeconomic barriers to the re-use of the waste streams from energy production from biomass. They confronted risk-based barriers with the opportunities of using the remains as soil fertilisers under clear standards of sustainability.

SfRA and SfSA are innovative by the exploration of alternative management solutions upfront in the risk- or sustainability assessment process. Perceived benefits are that assessors become actively involved in defining the key societal questions and in finding realistic approaches to minimize risk and optimize sustainability. As a result, the assessment results might be closer to (or better applicable for) management decision support. To our knowledge, practical applications of SfRA have not been published thus far. Given the utility-improvement argument that triggered the solution-focused approach, and the wider need for practical and societally important multi-metric sustainability evaluations such as those described by Riding et al. (2014), there is a momentum to re-consider some environmental problems currently addressed as single-metric risk problems, to redefine them in the wider context of sustainability assessments, and to apply the solution-focused paradigm for finding sustainable solutions to those problems. Therefore, the goal of this paper is to forward sustainability science by introducing a versatile and operational solution-focused approach for solving complex environmental problems, through the following activities:

1. defining Solution-focused Sustainability Assessment (SfSA) as the complement of SfRA;
2. providing practical though versatile guidance (both procedural as technical) for performing a SfSA for wicked problems;
3. illustrating the application of SfSA to solve a wicked environmental assessment and management problem: the disposal of slightly contaminated sediments from ditches in rural areas.

Note that the versatile framework is illustrated with one case study, and that this was done to show a practical application of the framework in solving a wicked problem. However, due to the highly variable nature of wicked environmental problems, the case study should not be interpreted as to be representative for all possible cases.

Table 2.1 (A) Features of wicked problems (Rittel and Webber (1973); Stahl and Cimorelli (2013)), and (B) strategies to cope with them (Roberts, 2000).

A) Wicked problem features	**Characteristics**
Multidimensional context	The problem context cannot be easily defined and agreed on by all stakeholders.
	The problem issue allows using and selecting multiple metrics.
Multiple stakeholder perspectives	The problem can be defined in many ways, including variation in the spatial and temporal scales.
	The problem requires complex judgments about the level of abstraction at which to define the problem.
Reflects non-optimality	There are no clear rules to finalize a multi-metric assessment.
	There is no single optimum.
Trade-offs among conflicting goals	There are better or worse conditions, not right or wrong ones.
Subjective, values-driven	There is no objective measure of success.
Learning-driven	Solving the problem requires iteration – every trial counts.
Stakeholder-driven learning	There are no given alternative solutions – these must be discovered.
Multidimensional legitimacy	The problem often has strong moral, political, or professional dimensions.
B) Coping strategies	**Characteristics and (dis)advantages**
Authoritative	A selection of people is asked to solve the problem; they get the necessary means.
	Advantage: reducing # of stakeholders reduces process complexity as (some) competing points of view are eliminated from the start. Disadvantage: relevant perspectives on problem and solutions may lack.
Competitive	Focus on contrasting points of view; best solution 'wins'.
	Advantage: weighing of wide variety of alternatives.
	Disadvantage: confrontational setting; discouraged knowledge sharing.
Collaborative	Engaging all stakeholders to find the most supported solution; Iterative exploration towards a common, agreed approach.
	Advantage: final output can be implemented with broad support.
	Disadvantage: process management complex; final management choice may be consensus based and sub-optimal from different stakeholder's perspectives.

2.2 Methods: defining Solution-focused Sustainability Assessment (SfSA)

2.2.1 Principles for SfSA

The SfSA framework has been designed to be specifically suitable to cope with wicked problems. It is therefore a versatile framework (Fig. 2.1), defined by six main steps, which can be followed sequentially but also – as needed given outcomes of the central communication interface step –iteratively. In essence, this implies that all steps have a key role in the process, co-determining the success of the process. It also implies that the pitfall of defining a fixed-process for a highly variable set of problems is avoided. Table 2.2 lists the basic principles for and characteristics of a SfSA approach and refers to the frameworks in which the basic principles have been operationalized earlier.

Table 2.2 Basic principles for and characteristics of Solution-focused Sustainability Assessment, and the frameworks in which these principles are operationalized.

Principles	Characteristics	Framework of origin
Solution focused	Initial exploration of multiple solutions. Compare options. Implement adaptive management. Consensus aim is concrete, stepwise, curing.	Solution focused Risk Assessment
Clear definitions on sustainability	Define the system boundaries (life cycle phases, spatial and temporal focus of the solution and its impacts), what to sustain (domains: people, planet, profit/prosperity (PPP)) and the basis for aggregation.	Sustainability assessments
Participative	Participation approach tailored to the problem. Wicked problems often require inclusion of many stakeholders and/or general public consultation.	Risk governance, Sustainability Assessments
Iterative	Finding solutions for wicked problems is seen as an iterative process, so that adaptive management is key part of the assessment process.	Adaptive management, Risk governance
Innovative	Innovative ideas (brainstorm) and solution-generation (with 'out of the box thinking') are essential. Also, exploring unrealistic solutions might be valuable, i.e. as a catalyser to find new 'out of the box' realistic solutions.	Solution focused Risk Assessment

2.2.2 Versatile and stepwise approach

We designed the stepwise Solution-focused Sustainability Assessment (SfSA) framework by combining the 'solution-focused' paradigm as designed in the risk assessment context with key elements of existing frameworks for especially risk assessment, sustainability assessment and risk management and -governance (Fig. 2.1).

The SfSA framework is primarily based on the merger of two classical frameworks, viz. those for risk assessment (U.S. NAS, 1983) and for risk governance (IRGC, 2008; Renn, 2008; Renn et al., 2011). The risk assessment framework has a standing tradition in environmental assessment and management. The risk governance framework expands on this, by including a pre-assessment step to acknowledge the existence of four types of risk problems, i.e., simple, complex, uncertain or ambiguous. The type of problem determines the assessment process and the stakeholder engagement. In both the risk assessment and the risk governance framework, the societal concerns are translated in risk information, which is then transferred to the risk management context.

These two frameworks were merged with the novel 'solution-focused' assessment paradigm originating from the risk assessment context (Finkel, 2011; U.S. NAS, 2009). Key to the approach is that, instead of mainly focusing on the problem, the focus is on solutions to the problem too. Both the problem and solutions are being explored upfront in the process, with minimal a priori limits to alternative solutions. A key element is allowing 'out of the box' enumeration of optional solutions for wicked problems, as those can offer pathways to innovative approaches when classical approaches – e.g., those focusing on single-metric risk assessments – have failed for the problem at hand. A recent example of such an expansion of the potential 'solution space' for a complex environmental problem is the focus on a smart spatial arrangement of chemical emission points (environmental arrangement) rather than on chemical management (chemical-oriented hazard evaluation, see Coppens et al. (2015).

Next, we added key elements of sustainability assessments to the SfSA design (Gorman et al., 2012; Sala et al., 2013; Zijp et al., 2015), especially the specification of varying societal views on the problem, on the definition of sustainability and on the decision context, followed by a transparent and participative translation of these views into methodological choices. Sustainability assessment is multi-faceted, as it encompasses 'people, planet, and profit' metrics, which will be further referred to as sustainability metrics. Participative translation refers to a process in which stakeholders are invited to discuss the design of the sustainability assessment. For example, considering the question which themes are important, or to which extent life-cycle thinking should be included in the assessment. Studies from the realm of Policy science have shown that transparent and participative processes enhance the trust in the outcomes and in the choices based on the outcomes (Lind et al., 1990; Lind and Tyler, 1988; Tyler and Lind, 1992; Van den Bos et al., 1988).

To cover the incremental learning aspect of dealing with wicked problems, characteristics of adaptive management approaches (e.g. Linkov et al. (2006), Kingsford et al. (2011) and Robinson and Levy (2011)) were incorporated in the SfSA framework. The success of the implementation of a chosen solution is evaluated iteratively, forming new SfSA cycles when needed. This evaluation again includes stakeholder participation, because the view on the degree of solving a problem depends on ones view on the problem (Gorman et al., 2012).

As stressed before, the framework requires various interactions across actors and assessment steps. Therefore, communication, as in risk governance and in Sustainability assessment, has a central position in the SfSA cycle, pertaining to communication across assessment steps as well as within and across stakeholder groups.

The process is described as six successive steps, but application of the framework allows iterations between steps. As an example, the quantification of sustainability metrics may yield novel insights that should be fed back into the step in which solutions are developed; or: the development of solutions can result in new insights for the Pre-assessment, e.g. on which stakeholders should be involved.

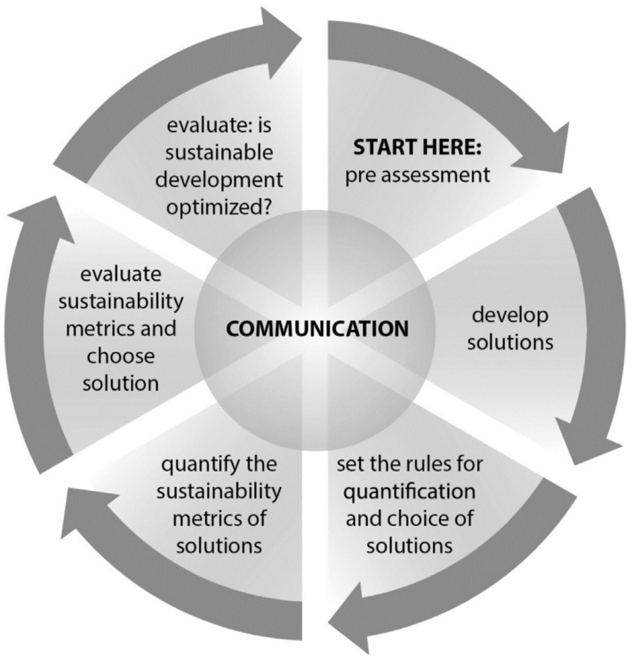


Fig. 2.1 Main process steps (clockwise) for Solution-focused Sustainability Assessment. Process steps representing the innovative solution-focused actions are marked with double outlines. Iteration across steps is possible. For details and guidance: see section 2.2.3, **Table 2.3**

2.2.3 Practical guidance

Practical guidance to support application of the SfSA framework is provided in this paragraph. Every case study asks for its own specific design based on the ingredients provided by the SfSA framework and thus the guidance provided here is not meant to be complete; it rather hints for implementation choices based on the authors own experiences. The case study (Section 2.3) provides a concrete description of applying SfSA.

In the first step, Pre-assessment, a literature and context review is performed to describe the problem, the context and limitations and opportunities from the viewpoint of different scientific disciplines (environmental science, social science, economics, etc.). A critical evaluation of the regulatory motives and hidden assumptions of the problem definition is part of this step. That is, a problem may be defined by a risk-based regulatory criterion that in itself is derived based on scientific insights and value judgement (e.g., the level of protection aimed at). Moreover, a stakeholder analysis is performed to gain insight in the different existing perspectives on the problem and the attitudes of the stakeholders towards sustainable development (Adams et al., 2015; Klewitz and Hansen, 2014; Zoeteman, 2012). Support for the identification of the relevant disciplines and stakeholders can be found in Hage and Leroy (2007), IRGC (2008), Reed (2008) and Pooley et al. (2013). Like in the example of Riding et al. (2015), this step serves to decide if the problem is indeed a wicked problem that asks for a move from a single-metric risk assessment perspective to a multi-metric sustainability assessment.

The second step, Develop solutions, focuses on finding possible solutions. For the identification of a wide array of solutions, brainstorm techniques can be applied, such as 'sustainability jam sessions' (Carlsson et al., 2015). Subsequently, solutions are selected based on the knowledge and preferences of the participating stakeholders (including the scientists). This selection is both a sanity check (which solutions potentially make sense from different perspectives) as well as a 'trading zone' for stakeholders with different perspectives (Gorman et al., 2012). Gorman et al. defined a trading zone as a setting in which ideas, resources, and solutions can be exchanged across different communities of interest, with as goal to cooperate in finding a solution that addresses the concerns and potential requisites of all parties. With other words, stakeholders with different perspectives on the problem and on the goal of an assessment cooperate to derive mutually agreed solutions through discussion and learning.

Linked to the first steps (Pre-assessment and Develop solutions), starting and running the stakeholder involvement and managing the stakeholder contributions can be very challenging. It is commonly not easy to get all relevant stakeholders involved, and it is difficult to retrieve input from large groups of people. Furthermore, it is challenging to merge the diversity of perspectives and knowledge into manageable input for the next step. Getting involvement of stakeholders requires clear communication on the importance of their role in the process as a whole and the benefits of their participation. When input is required from larger groups of stakeholders, working in subgroups can be an option to allow all stakeholders to ventilate their knowledge and perspective. Various tools and approaches exist that can support these processes such as hexagon brainstorming, GroupSystems and Group Decision Rooms (George et al., 1992; Rouwette et al., 2002). These tools can also support transparent ways to draw conclusions from the participation process. Merging the diversity of inputs can partly be done during the meetings by the participants themselves. The other part is clear communication on the conclusions drawn and, when necessary, iterations of the process: another meeting in which the results are discussed and adjusted until the solution scenarios defined reflect the perspectives of all stakeholders as much as possible, while potentially contributing to improved sustainability of the situation.

Regarding the definition of the assessment problem, it is key to not only delineate the problem (Pre-assessment) and to develop solutions, but also to define sustainability in the context of the case study. Allen et al. (2003) suggest to pose four questions to be answered in order to define sustainability: 1) what is to be sustained; 2) for whom is it being sustained; 3) for how long is it being sustained and 4) at what cost is it being sustained? With other words: the system boundaries (the spatial scale, the temporal scale and the life cycle stages) and the sustainability themes (the social, environmental and economic aspects) are to be transparently selected before an assessment method can be proposed (Zijp et al., 2015). Therefore, the following step, Set the rules for quantification and choice of solutions, is the step in which the problem definition is discussed and analyzed in such a way that appropriate methods and metrics are chosen to describe and to quantify the current situation and the optional solutions. The assessment does not require to include all aspects of sustainability, but to determine which are relevant for the case study (Laniak et al., 2013). Selection of system boundaries and sustainability themes is performed together with the stakeholders. The next step is to select the appropriate methods to quantify the themes considered to represent sustainability. In complex situations this is often an iterative process between the group of stakeholders that agree on a conceptual model on what the assessment should look like and the assessment experts that translate this conceptual model into an operational model. A satisfying translation is not always possible due to a lack of knowledge methods and/or data. Therefore, the design of the operational model might require interaction between the stakeholders and the experts in order to make the design as optimal as possible. Transparent translation of the conceptual model into the operational model results in support for the assessment (Lind et al., 1990; Lind and Tyler, 1988; Van den Bos et al., 1988). In addition, the options for presenting the results of the chosen metrics are explored before the actual assessment takes place. The presented results should not only be salient, credible and legitimate (McNie, 2006), but also useful (Laniak et al., 2013) in the sense that they are understood and ready for use by the decision makers (Brewer and Stern, 2005). In order to present results of complex projects in an understandable and meaningful way, they can for example be collated into aggregated scores (Özdemir et al., 2011). Aggregation however has as disadvantage that it is value-laden (Gasparatos and Scolobig, 2012), it can add to the uncertainty of the results and information is lost (Özdemir et al., 2011). Another option is to present the results as a database with an interface for the decision makers with which they can explore the results interactively (Laniak et al., 2013). Finally, in this step, the procedure on how to decide on which solutions are to be implemented is delineated. This procedure is more or less comparable with the procedures of Multi Criteria Analysis, as otherwise transparency cannot be guaranteed (Dodgson et al., 2009; Linkov et al., 2006). It is not always possible to quantify all themes. In such cases, rather than neglect important themes and metrics, there are options that allow for inclusion of qualitative indicators, such as the REGIME method (Hinlopen and Nijkamp, 1990).

As fourth step, Quantify the sustainability metrics of solutions, provides quantitative information on the chosen metrics according to the appropriate methods, and optionally provides a collation of the multi-metric outcomes of the current problem situation and its alternative solutions. Scientists are the main actors in this step, operating in a mainly scientific-technical context of the agreed technical approaches. The approaches can range from monitoring and modeling to surveys and expert judgement. Depending on the case study and the number of procedural iterations that are accounted for, this step can serve various roles with various requirements. For example, a quick scan quantification step can support the selection of optional solution directions (support step 2 iteratively) or pre-assess the relevance of themes (support step 3). However, all these assessments, either thorough or quick, should be based on the scope set in step 3 by the stakeholders.

SfSA is designed such that decisions on the procedure and metrics to be applied for evaluating solutions are made before the actual quantification of the metrics takes place. Thus, in the step, Evaluate sustainability metrics and choose solution(s), the agreed technical and decision procedures (step 3) are applied. Optional techniques for interpretation, harmonization and weighting of metrics are many. See also the remarks above. Examples are provided by Huang et al. (2011) and Bruggemann and Carlsen (2012). The pre-selected decision procedures describing the role of the different stakeholders in the decision process are applied, eventually yielding information on the ranking of solutions according to the agreed metrics, or an agreed aggregated metric. In an ideal case, the 'best solution' is defined by all sustainability metrics being optimal, so that this 'best solution' can be implemented. However, for most wicked problems, the solutions imply trade-offs on sustainability metrics, so that there is no single optimum. A management choice is thereafter made, by applying the agreed output interpretation approach.

Finally, the step 'Evaluate: is sustainable development optimized?' is the evaluation of the process and its results. After implementing the selected management option, it is checked if changes in metrics due to the implementation of an agreed solution indeed show (changes in the direction of) sustainable development metrics according to expectations. The monitoring and evaluation may also unveil novel conditions, or novel solutions supporting further optimization. When this occurs, the SfSA-cycle can be repeated for further sustainable development.

It should be noted that the solution-focused paradigm does not substantially increase project turnover time. The application of the framework implies making a comparative assessment between the current situation and alternative solutions. The project turnover time is typical for each problem definition, and the extension in project time is only limited, and mainly related to the pre-assessment and solution-development steps. There, the problem description and the optional solution scenarios are defined. Whilst this may expand project duration, it also may reduce the time required in finding a suitable and supported solution to the problem. After the first round of comparative assessment, the framework can result in the first implementation of the selected solution. Table 2.3 provides guidance per step of the SfSA framework. Each step is characterized by various checkpoints, being optional points of attention or actions to be considered.

Table 2.3 Guidance of optional checkpoints per step of the SfSA framework.

Optional checkpoints	Suggested methods
Pre-assessment	
Assemble a multi- disciplinary team	Determine the disciplines that are required to guide the process and to perform the different steps. The role of the team members is that they deliver continuous input from different disciplines in the whole process. Next to the team, (other) experts can provide input at several stages in the process. The team may e.g. consist of experts in social interactions, environmental impacts, economics and jurisdiction; but also a communication expert and skilled process manager should be part of the team from the start (and not only on the moments that communication or a stakeholder meeting is planned).
Execute a stakeholder analysis	There are several ways to perform a stakeholder analysis. See: http://www.irgc.org/risk-governance/stakeholder-engagement-guide/. Bottom line is that all relevant perspectives are included. Be transparent on which stakeholders should be included or excluded in the project. In reality, the stakeholder analysis continues during the whole project, e.g. the selection of solutions can raise new insight on stakeholders to be involved.
Formulate the problem definition(s) and views on sustainability from different perspectives	Stakeholders' view on the problem and the definition of what sustainability can be assessed using the Sustainability Assessment Identification Key (SAIK) (Zijp et al., 2015). The goal is not one definition, but an inventory of the different views. Agreement on solutions can overarch different views on the problem and its context.
Develop solutions	
Together with stakeholders develop a list of possible solutions and perform a reality check in order to select the most feasible ones.	Every group of stakeholders requires its own participation process. Examples of how to organize this participation can be found here: http://www.irgc.org/risk-governance/stakeholder-engagement-guide/.
Set the rules for quantification and choice of solutions	
Decide about metrics to be used to quantify or qualify sustainability and decide about if and how different indicators should be aggregated	The SAIK (see Pre-assessment) can be used to provide the conditions the quantification should encompass (which system boundaries, which themes, the necessity of quantitative aggregation) and translate these to method/indicator selection. Differences in the view on how to deal with methodological choices can be solved by implementing them as sensitivity scenarios, e.g. such as provided in the Life Cycle Impact Assessment method ReCiPe (PRé et al., 2013).
Draft the decision procedure.	Be transparent about how the choice for a solution or a set of solutions will be made: who will make the decision and how do the different stakeholders participate in this decision? A Multi Criteria Analysis (MCA) can support this process. In a MCA the results of the quantification are made comparable (e.g. standardized or normalized) and are then weighted. The weights can be determined in a participative process. An MCA, also the standardizing and normalisation, are value-laden and methods should be selected transparently and with care.
Quantify the sustainability metrics of solutions	
Apply the selected approaches to quantify, scale or qualify metrics	Various overviews of methods exist, e.g. De Ridder et al. (2007); Ness et al. (2007); Wrisberg et al. (2000). Applying methods and finding data will often results in methodological choices that require feedback to the stakeholders (iteration with previous step).
Evaluate sustainability metrics and choose solution(s)	
Follow the decision procedure set in the previous steps. Thoughtful presentation of the results.	The results of the previous step should be presented in a way that supports the decision process. Due to complexity of the case study, this can be challenging. Next to graphs and tables that summarize the results, an option is to build a database with the results and provide it with an interactive interface such that visualisations can be made on request or even by the stakeholders themselves (Laniak et al., 2013).
Evaluate: is sustainable development optimized?	
Evaluate how the process influenced the problem: collect perspectives of stakeholders and monitoring data on (key) metrics	Organize a meeting with all the stakeholders in which the process and it outcomes are discussed. Which problems are not solved and require a new or revised project?

2.3 Results: SfSA application in a case study

2.3.1 Introduction to the case study

The SfSA framework was applied retrospectively to a case study on the management of contaminated sediments. That is, when considered in hindsight, the case study concerns a wicked environmental problem (see Table 2.4), in which the mono-metric contaminant risk focus has changed into a multi-metric sustainable sediment management focus, while the problem-solving process showed all elements now collated in the SfSA framework. The case study relates to the management of slightly contaminated sediments in ditches in rural areas, worked out for the Netherlands (though similar problems occur in many landscapes around the globe).Like the study of Riding et al. (2015), it appeared that the current system of judging sediment misjudged (over-estimated) local risks of sediment deposition on land, and that the problem description, analysis and solution asked for wider views.

Sediments are continuously formed in waterways (here: ditches). Preferably, sediments are deposited on adjacent agricultural land (Fig. 2.2). For water management, the goals of that practice are to maintain drainage capacity and avoid flooding, and for soil management to improve soil fertility and to counteract soil-subsidence. The classical approach to handle uncontaminated sediments consisted of local re-use and no off-site treatment costs. However, this practice became suspect since the 1970s due to sediment contamination. Sediment management decisions were thereupon facilitated by defining national sediment contamination classes, from Class 0 (clean) to Class 4 (heavily contaminated). The sediment management policy was focused on avoiding and reducing contaminant risks for land soils. That is, Classes 3 and 4 were not allowed to be spread on land, and deposition of Class 2 sediments would be phased out in a decade (VROM, 1993). The scientific basis for that policy was a combination of i) generic environmental quality criteria for chemicals, related to protection and sanitation of contaminant concentrations in soils and ii) an evaluation of expected aggregation of hazards due to the presence of mixtures of contaminants. In turn, the basis for those soil quality criteria dates back to the 1980s and 1990s. They were derived using No Observed Effect Concentrations (NOECs) which were established in laboratory tests in which soil inhabiting species were exposed to individual chemicals.

Fig. 2.2 Removal of sediments from waterways in the rural areas of the Netherlands and spreading of the material on adjacent land.

Sediment management practices developed into a wicked problem (see Table 2.4) during the application of the policies in the 1980s and the 1990s. First, it appeared that large volumes of sediment were characterized as Class 2 or higher, whilst the expected reduction of sediment volumes of these classes was not occurring. The perceived hazards of all those sediment loads resulted in a halt to the sediment deposition, increasing backlogs of sediments in ditches, and increasing probabilities of flooding. Second, stakeholders increasingly expressed different views on the problem and its solutions. One: water quantity management highlighted flooding problems, and stressed sediment removal. Two: water quality management focused on emission reduction needs and removal of contaminated sediment. Three: soil and groundwater quality management focused on protecting soil quality, highlighting the maintained use of the soil-protecting Classification scheme. Four: water boards, as competent authorities and sediment management practitioners, needed to consider all these matters, including both budgetary costs of sediment management and realistic management options. Finally, like in Riding et al. (2015), public perception was on perceived pollution (human and environmental) risks of deposition and on water board taxes, and not on the wider scope of sustainable sediment management. The increasing 'wickedness' and urgency to break the decision inertia given increased flood risks was a national problem of 'large numbers'. Regular (e.g., once per three or five years) sediment removal is needed for 330,000 km of Dutch ditches (CBS et al., 2015). The removal of the backlog of 200 million ton of sediment of varying quality in the period 2002-2011 would require a budget of 975 million €, and substantial yearly maintenance costs thereafter (AKWA, 2001). Thus, sediment disposal evolved into a wicked environmental management problem, triggered by the following metrics:

- human health risks that could occur via contaminant transfer from sediment to soil to food resources to humans;
- ecological risks: threats to ecosystem integrity; and
- socio-economic impacts: the costs of sediment transport and optional off-site treatment as well as of flooding and soil subsidence. Flooding has serious and acute consequences for society, as exemplified for e.g. England and Wales (Wheater and Evans, 2009).

The SfSA case study lasted for less than a year for the research definition and execution part (steps 1 - 5). However, before that, the analysis of the management problem and optional solutions had already taken more than a decade, triggered by the lack of volume reduction of (especially) Class-2 sediments. After the research project, the decision to change sediment management policies took approximate one other year, in which the chosen solution was translated into policy principles and sediment quality judgment rules. That process included synchronization with changes in other environmental management fields, especially soil contamination management.
More information on the evolution of the problem and its regulatory context is described in the Annex (§2.7).

2.3.2 Evaluation of the SfSA framework

The basis for finding a way to deal with the sediment removal (ditch) and disposal (land) problem is described below, referring to the steps of the SfSA-framework. Note that communication was a continuous and important process throughout the project, from the Pre-assessment to implementation of the resulting novel sediment policy.

2.3.2.1 Pre-assessment

First, an inventory was performed on the type and magnitude of the sediment management problem, including aspects of costs and options to manage (slightly) contaminated sediments (AKWA, 2001). The inventory showed the societal urgency of solving the problem, and aspects of risks and costs involved. In view of scientific developments, the government concluded in addition that the sediment classification scheme was outdated and didn't serve its purposes (cost-)effectively (VROM, 2003).

Table 2.4 The wicked problem features of the case study.

Feature	Case study
Multidimensional context	The early mono-metric chemical risk perspective prevailed for a long time, triggered by initial sediment contamination observations. A clear multi-metric perspective emerged when the contamination classification policies triggered other risks and/or large costs. Management of the contaminated sediments appeared to involve a trade-off between potential human health impacts, various ecological impacts and various socio-economic impacts.
Multiple stakeholder perspectives	Waterboards (water quality and quantity responsibilities), provinces and municipalities (water quality, soil and groundwater responsibilities), the national government (policy responsibility), farmers (food quality responsibility) and the public (waterboard taxes) all had a different perspective on the problem and on the system-boundaries within which an alternative management solution could be found.
Reflects non-optimality/ Trade-offs among conflicting goals	Thinkable solutions all imply trade-offs between different dimensions (human health, ecological and socio-economic), e.g., non-removal ➔ floods; out-of-system removal ➔ soil subsidence and high costs; et cetera.
Subjective, values-driven	There is no objective measure of success, as 'no-spread related to soil quality protection' is successful when seen from the soil quality point of view, but not from the flooding point of view.
Learning-driven/ Stakeholder-driven learning	Finding a solution-scenario involved various iterations in which the problem (pre-assessment), possible solutions (develop solutions), the requirements for the assessment method (rules for quantification) and the quantifications itself where discussed with the stakeholders.
Multidimensional legitimacy	Finding solutions involved discussions on, for example, the definition of the standstill principle and the precautionary principle.

Mandated by a national policy initiative, a multi-disciplinary taskforce was formed. It consisted of representatives from policy, science, sediment management practice and innovation management, with an independent chair. Its task was to develop the outlines and requisites for a new sediment management policy. A stakeholder analysis was performed by mapping the different stakeholders and their roles within the context of the case study. Most stakeholders could be identified given their existing activities and involvement in discussion groups on policies and practices regarding soil quality protection, sediment management or water quality and quantity management. This inventory resulted in a wide representation of stakeholder views, as it implied involvement of national, regional and local authorities, farmer organizations, non-governmental organizations, et cetera. Additional stakeholders were found via snowball sampling (Hage and Leroy, 2007). All located stakeholders were invited to discuss their concerns and views to the problem in a national workshop. Furthermore, a communication strategy was designed, including a regular newsletter for stakeholders and information for the general public (e.g. Ministry VROM et al. (2003)). Both risks and benefits of sediment applications were described in an informative 'honest broker' style, supported by schematic line drawings (Doelman 2003). This communication resulted in recognition of both the advantages as disadvantages of sediment re-use and disposal instead of the prevailing 'waste only' perspective (like in the example of Riding et al. (2015)).

2.3.2.2 Develop solutions

The preferred solution agreed by the stakeholders was formulated as: to deposit the sediment on adjacent land as long as it is sustainable and safe, with an emphasis on local re-use (deposition) when possible. If not, the sediment could be used in sealed infrastructural applications or be treated. The central question following from these preferred solutions was: 'when is deposition sustainable and safe?' To make the formulated solution operational, the taskforce formulated draft requisites for a novel policy, and evaluated them with the stakeholders. The novel policy should:

- protect man, the quality of agricultural products and the environment;
- align with sediment, water and soil contamination policies, e.g. overarching European regulations;
- be simple and transparent;
- be easy to control during environmental inspections;
- encompass a practice-oriented decision-support system;
- be implemented by regional competent authorities for the areas under their supervision; and
- focus at opportunities, not merely on risks.

To reach these goals it was acknowledged that there was a need to broaden the decision basis of sediment management in three ways:

1. change from a hazard assessment basis to a risk assessment basis, by including characteristics of the receiving system (e.g. open field application compared to sealed application imply different risks, as exposure between these options differs);
2. change from a compound focus to a system-wide focus by taking into account the spatial-temporal behaviour of the compounds, and the mixture of compounds, in the sediment and the receiving soil (Posthuma et al., 2008); and
3. instead of an easy to apply and soil-protection oriented 'one out all out' principle (as a characteristics of the Classification system), explicitly weigh the different metrics (human health, ecosystem health, product safety and costs) locally for decisions on area-specific sediment management.

2.3.2.3 Set the rules for quantification and choice of solutions

In this step, the required procedural and technical choices were made on the quantification of sustainability metrics, and the eventual evaluation of the possible solution scenarios. The research problem was articulated by defining the system boundaries. Spatially, all sources, fates and sinks of chemicals in the water-sediment-land system were considered on a regional basis, linked to the organisational level of the water boards. Temporally, the present situation was taken as reference point for the modelling. These system boundaries were input for a new scientific modelling tool for sediment risk quantification. The tool should be able to quantify the impacts of alternative solution scenarios on the themes human health, ecological health, agricultural products and costs. The metrics are further described in section 2.3.2.4.

Procedural, it was agreed that the quantification of scenario results was designed to be strictly separated from policy or value judgment. That is, all quantitative 'raw' research outputs would be the subject of an iterative evaluation process, in which the values of all metrics for the alternative solution scenarios would be considered and weighted by the stakeholders. This could imply iteration between the quantification and decision steps, given the role of policy makers (deciding on risk management criteria). As the results generated by the tool would concern multi-metric outcomes for the many thousands of sediment qualities available, it was agreed that those multi-metric results would be discussed with the stakeholders in the format of summary outputs, such as box-and-whiskers plots. Such plots would summarize the ranges of effects of sediment deposition on the selected metrics under an envisaged decision scenario, and would allow easy comparison among thinkable management scenarios. The case study thus reflects an example of an evaluation of alternative policy decision options based on a large database of sustainability metrics.

2.3.2.4 Quantify the sustainability metrics of solutions

In this step, the system boundaries and themes selected were translated into indicators. A model tool was made operational to quantify the indicators. Data was either available (AKWA, 2001) or estimated (the costs). The model tool consisted of (1) an exposure assessment model, which enabled prediction of contaminant concentrations in soils over time, given a sediment deposition scenario and sediment contamination levels, (2) effects assessment models to quantify expected impacts on human health, ecosystems and agricultural product quality, and (3) the option to implement different decision criteria for the different metrics (both exposure- and risk based options could be applied). The application of the tool to the field data set resulted in output metrics for individual sediment management scenarios, as well as in summary boxplots for the set of sediment loads. A selection of the results from this step is presented below. The selected results illustrate the types of results in the type of formats that were communicated with the stakeholders. Also intermediate results were presented and discussed to explain all (scientific) steps made, to promote understanding of the results, and understanding of why those results differ from the well-known sediment Classification schemes. The principles, modelling steps and detailed outcomes of the analyses of exposures, expected risks, and the results under alternative management scenarios are described in detail in (Posthuma et al., 2006a), (Van Noort et al., 2006) and Posthuma et al. (2006b).

The inventory database and spatial variability of the problem. A spatial map derived from the AKWA-database shows that sediment contamination problems vary between regions, both in volume and contamination-Class (Fig. 2.3, left). This is partly related to the spatial variability of the background concentrations of contaminants in soils (Fig. 2.3, right). This Figure shows the mixture toxic pressure of current contaminant levels in soils, on soil ecosystems (mixture toxic pressure explained below). The spatial differences were expected from- and confirmed by various other studies on soil chemical loads (Harmsen, 2004; Harmsen et al., 2012; Spijker et al., 2011). Policy makers and stakeholders acknowledged these spatially explicit and variable results. That paved the way for allowing regionally optimized approaches in sediment management.

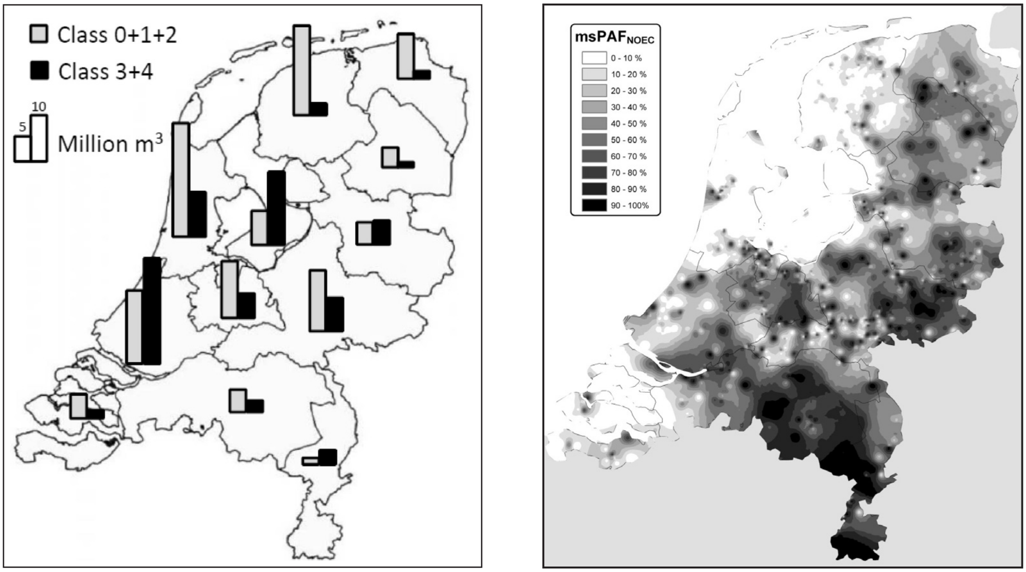


Fig. 2.3 Left: contaminated sediment volumes and the classical sediment quality Classification for various regions. Right: spatial variability of mixture toxic pressure ($msPAF_{NOEC}$) of background levels of contaminants in soils in the rural areas of the Netherlands (for explanation: see text).

Metric 1: Chemical fate of the contaminants in soil. The fate of the contaminants in soil determines the impacts on human health, the ecosystems and the quality of agricultural products. Moreover, changes in concentrations can be judged against sustainability principles like standstill on the short or long term, and against policy decisions such as allowance of transgression of generic soil quality standards for a short term.
The concentration metrics varied widely across areas, substances and scenarios. The scenarios were based on the initial spatial variability in contamination, different sediment deposition regimes (e.g., once per five years), and the physico-chemical behaviour of compounds in the environment, such as degradation of organic compounds. As an example, the concentration change patterns for benz[a]anthracene (B(a)A) and copper over time (Fig. 2.4) show how concentration patterns differ for different chemicals under the same deposition scenario. Most scenarios resulted in a long-term equilibrium soil concentration (much) lower than the original sediment concentrations. In a few scenarios, disposal of sediments resulted in concentration reduction in soils due to 'dilution' in the soil-sediment mixture (the local sediment was less contaminated than the soil).

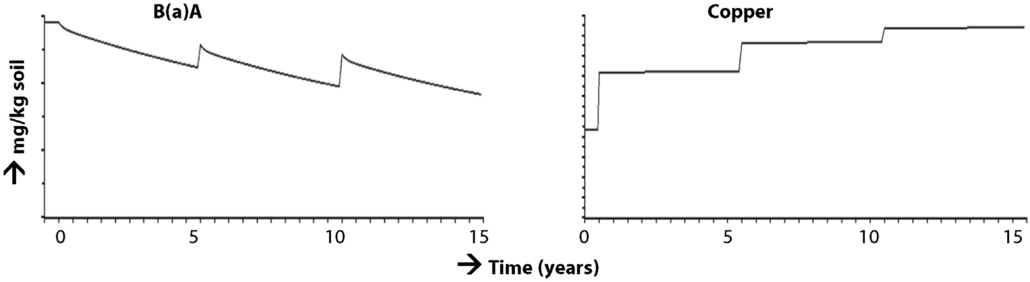


Fig. 2.4 Soil concentration change patterns for benz[a]anthracene (B(a)A) and copper for one of the sediment deposition scenarios (one sediment sample, 3 applications, 5-year application interval).

Metric 2 and 3: Human health risks and risks for agricultural product quality. In Fig. 2.5 the human health-based risk characterization is presented as a function of time. The saw-tooth curve represents the predicted lifelong-averaged human exposure (mg•$kg_{body\ weight}^{-1}$•day^{-1}) resulting from the CSOIL exposure model (Brand et al., 2007; Swartjes, 1999). The Figure includes three options to judge the health implications of the expected exposure (horizontal lines). The first is the exposure associated to the regulatory-defined maximal permissible risk (MPR) for human health impacts. The second is the negligible risk (NR), defined as a fraction (1/100th) of the MPR. The third is a land-use specific exposure criterion. That is, exposure criteria can be judged dependent on land use, as local human exposure not only depends on soil concentrations, but also on the 'intensity' of human land use; e.g. food production implies a higher exposure than residential use, at the same

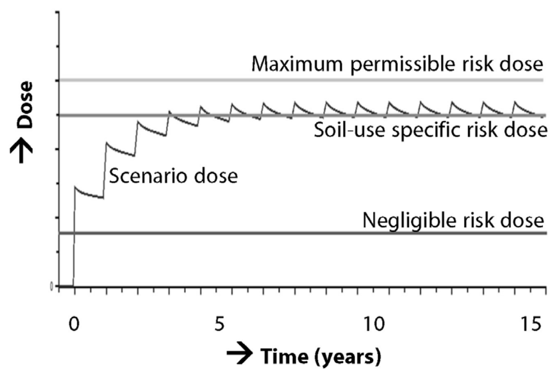


Fig. 2.5 Human health-based risk characterization as a function of time, given three optional policy criteria to judge exposure risks.

concentration. The Figure illustrates that different policy choices would imply different fractions of sediment to be deposited on land: by choosing MPR as generic human health safety criterion, more sediments can be deposited on land than by choosing NR. When land-use is taken into account, sediment spreading on land is decreasing from MPR>Land-use specific>NR.

The judgment of agriculture product quality follows the same principles, whereby soil concentrations are translated into product concentrations, which are subsequently judged against stringent product quality criteria.

Metric 4: Mixture risk for soil ecosystems. Fig. 2.6 summarizes the results for the metric 'mixture risks for soil ecosystems'. The Y-axis here represents the multi-substance Potentially Affected Fraction of species exposed beyond their No Observed Effect Concentration, which is a mixture risk metric (Posthuma et al., 2002). This metric expresses the fraction of species expected to be exposed beyond their no-effect level at a given soil mixture exposure. An increase of this metric quantitatively relates to an increase of mixture impacts (Posthuma and de Zwart, 2012). Despite the scientific weakness of the use of NOECs to derive this metric (Landis and Chapman, 2011)), it was key to apply this metric in the context of communication with the stakeholders. Most were familiar with the concept of protecting ecosystem integrity (the protection endpoint) via the so-called '95%-protection level', at which 95% of the species would be fully protected, i.e., exposed at a level not inducing impacts. In Dutch environmental law, the Maximum Permissible Risk (MPR) level for ecosystems has e.g. been defined via the 95%-protection level. Fig. 2.3 (right) already presented the spatial differentiation of this metric regarding existing soil contamination, prior to sediment deposition in the solution scenarios.

The mixture toxic pressure for soil ecosystems increases with increasing contamination Class (Fig. 2.6), suggesting that the original classification system was on average correctly indicating gross risk increase with increasing Class number. However, the Figure also shows that the classification system is over-protective

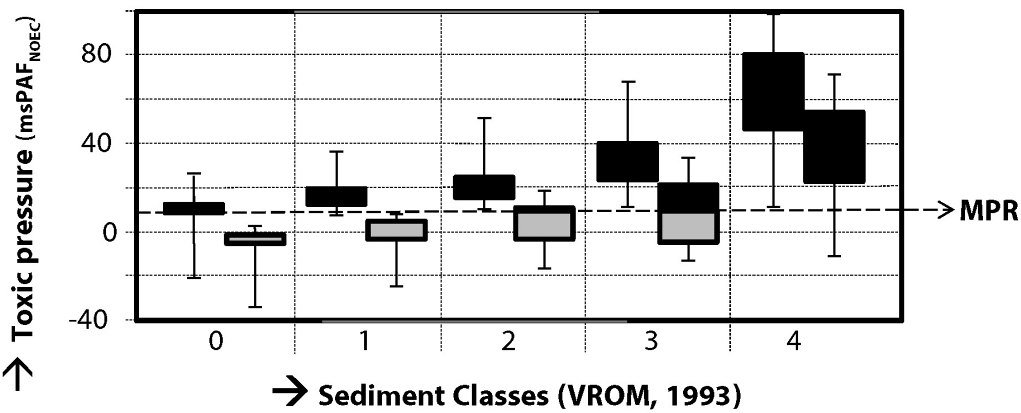


Fig. 2.6 Box and whiskers plots of approx. 15,000 data points (cases) of disposal of sediment on land. Each case was assigned first to the classical sediment Classification system (X-axis classes). The Y-axis is the (change of) mixture toxic pressure (fraction of species exposed beyond their no-effect level msPAF$_{NOEC}$). Within each of the Classes the left bars are absolute mixture toxic pressures of soils after sediment deposition and the right bars are mixture toxic pressure changes of soils (toxic pressure difference between the situations before and after sediment deposition). Each bar is colored according to an optional maximum risk criterion (to be chosen in the next SfSA step); here MPR (Maximum Permissible Risk, MPR, equal to a mixture toxic pressure of 5% at NOEC-level). Black bars: transgression of MPR due to deposition would result in 'no spread', grey bars: no transgression of MPR would result in 'spread'.

for a large fraction of sediments, and that Class-2 sediments in some cases appeared to imply higher risks than some Class-4 sediments. Thus, for individual cases, the classical system misjudges local risks. The novel metric provides a quantitative insight into risks for soil ecosystems that is different from the Class-system. The difference is caused by accounting for processes that occur in the system, like breakdown and bioavailability modification in soil systems (usually lowering risks) and mixture effects (net risks of a mixture).

When background concentrations in soils are taken into account (right bars per Class in Fig. 2.6), the increases in mixture toxic pressure in soils were less than when assessed without background concentrations, indicating that changes in ecosystem risks were lower than the predicted absolute ecosystem risks. The change may even be a reduced risk, when the background soil contamination is higher than that of the deposited sediments (values Y<0).

To interpret these outcomes, an optional policy decision boundary is added, in this case related to the MPR for soil contaminants (Y= $msPAF_{NOEC}$ = MPR = 5% exposed below their NOEC = 95% of the species assumed in policy to be fully protected). This results in an indication of the fraction of sediment volumes that cannot (black) or can (grey) be spread when making policy choices (next step) on inclusion of background concentrations and MPR.

Metric 5: costs. The evaluation of alternative solution scenarios concerned mostly the urgency of implementing a novel sediment policy triggered by increasing flooding risks and the evaluation of risk information for the endpoint human health, agriculture product quality and ecosystem integrity. However, alternative sediment management scenarios could also be evaluated regarding cost aspects, whereby the lower and upper cost boundaries are defined by the extreme scenarios: spread all or none of the sediment volumes locally. The cost difference between these extremes is estimated at 32 Euro per m3 (Rijnland, 2007). Given the inventory data on sediment volumes in ditches (AKWA, 2001) one illustration of the cost difference for different solutions is provided, to show the magnitude of cost differences at stake. In the example, the possibility to allow local spread of Class-2 sediment to MPR levels instead of the planned no-spread policy would reduce the costs of sediment management from 1.6 billion to 320 million Euros. A further exploration of costs under different policy choices showed that alternative policy choices would have large effects on sediment management costs and an unequal cost distribution across water boards. The less stringent the judgment criteria of spreading contaminated sediment, the lower the cost of sediment management, but the higher the potential costs of soil contamination management.

2.3.2.5 Evaluate sustainability metrics and choose solution

The group of stakeholders explored the results of all metrics generated in the previous step in an iterative way. The researchers provided insight in the base line scenario (the initial policy to deal with sediments) and in the scientific outcomes of the optional new scenarios. As a result, the views of experts as well as of non-professionals changed from 'almost always suspect sediments' to an assessment in the tradition of Paracelsus: not the compound (or the Class name), but expected quantitative risk estimates were interpreted to imply risks and impacts. This allowed for a nuanced evaluation of advantages and disadvantages of decision scenarios. Eventually, this resulted in the formulation and adoption of a novel sediment assessment and management policy in 2008 (Osté et al., 2008; Wezenbeek et al., 2007). Making use of all generated insights into systems-level changes of the selected risk and sustainability metrics, the key choices made prior to implementing the new policies were: (a) the national volume of sediment being spread should remain the same, (b) practical judgments would not take into account background soil concentrations and would thus not allow to consider risk change (see e.g. Fig 2.6), and (c) the mixture toxic pressure on ecosystems was separately judged for metals and organic compounds. The latter relates to vastly different environmental fates and expected impacts of metals and organic compounds.

2.3.2.6 Evaluate: is sustainable development optimized?

The novel policy was implemented, and evaluated in two ways after some years. First, Gadella (2011) reported about implementation practices, and mentioned some latitude for further sediment management innovations. It was suggested to consider a broader set of chemicals, according to the worldwide concerns about emerging chemicals and regional concerns on some specific compounds. Second, Harmsen et al. (2012) evaluated whether model predictions were confirmed by field data, such as a low or negligible increase of soil concentrations of contaminants upon long-term spreading of sediments. The study showed that the model results were generally conservative as compared to the field data, and that the novel policies do neither compromise soil quality nor increase risks. Sediment management practitioners discussed the novel policy and found that it allowed for local optimization of sediment management (e.g., Nieuwenhuis and Ellen (2007)), using the sustainability principles and metrics with the local conditions as input.

Though not considered in the case study, it can be noted that sustainable sediment management may be further improved by including other metrics, amongst others emissions of greenhouse gases (emitted as a consequence of potential sediment transport and off-site treatments), physical hazards related to transport of sediment, or the degree of soil subsidence, and additional solution-scenarios in which sediment is not considered waste but rather a useful resource for various purposes (like in Riding et al. (2015)).

2.3.2.7 The case study and the SfSA principles

The development of the new policy for sediment management followed the process steps (Fig. 2.1) and principles (Table 2.2) of SfSA, including the evaluation of success and exploration of adaptive management options as suggested. The process was solution- rather than problem focused. Sustainable development was not an explicitly mentioned goal of the study itself. However, the system boundaries and endpoints were clearly defined together with the stakeholders and the endpoints were chosen in both the People, Planet as Prosperity domains of sustainability. The process was highly participative and iterative and resulted in both an innovative assessment tool (a quantitative way to characterize expected impacts) and in new policy. The Annex (§2.7) provides an analysis on the suite of earlier policy- and research steps made since sediment management evolved into a wicked problem. It shows how various principles and process steps of SfSA have historically been increasingly applied in attempts to deal with the problem. The historical process of increasing use of SfSA steps suggests that the application of all steps of the SfSA framework seem to be needed for a successful process of solving the wicked environmental problem of sediment management.

2.4 Discussion

This paper introduced SfSA as a logical expansion of SfRA. We motivated why and how the SfSA framework can be applied in addressing wicked environmental management problems, and propose versatile guidance to support its use. The utility and implementation success of the approach is illustrated with a case study.

The strengths and weaknesses of the solution-focused approach have been discussed in the literature (Davies, 2011; Finkel, 2011; Hope, 2011; Paustenbach, 2011; Rodricks and Levy, 2013). Confronted with the case study experiences, there is a net positive match between perceived and experienced advantages, and practical solutions were generated for perceived disadvantages (Table 2.5). For example, the risk of false certainty on risks, a disadvantage according to Paustenbach (2011), Hope (2011) and Rodricks and Levy (2013), was addressed in the case study by applying conservative modelling principles (over-estimate concentrations and risks in soils by selecting the most conservative data on breakdown rates for organic contaminants) to account for uncertainties in the various modelling approaches, and by validation studies.

Table 2.5 Advantages and disadvantages of the 'solution-focus paradigm', according to various authors (left) and from our case study experience (right). Citations: 1 = Finkel et al., 2011; 2 = Davies, 2011; 3 = Paustenbach, 2011; 4 = Hope, 2011; 5 = Rodricks and Levy, 2013.

Literature	Ref.	Experience from the case study
Advantages:		
The assessment contributes to decision-making and thus to sustainable development ('opportunities'), instead of being restrictive ('limitations'), which is often the case with risk assessment.	1, 2	Chemical risk assessment (basic to the sediment class system) was put in the broader perspective of finding a sustainable multi-metric solution, given urgent societal needs. Opportunities were defined, evaluated and communicated.
Stakeholder support of the final solution direction; alignments of focus and interests of different experts and stakeholders	1, 2	There is indeed support for the solution and the research to underpin it; details of the solution are still under discussion, in line with the adaptive management approach.
Rich information basis and complete analysis, including risks, uncertainty, and the costs of control and juridical aspects.	1, 2	The approach led to input of many parties from different perspectives, resulting in a suite of sustainability metrics, and consideration of initially opposing policy viewpoints.
Disadvantages:		
It creates false certainty on the applicability of the outcomes of the analyses compared to other approaches. It does not solve that data and models to quantify risk scenarios might be highly uncertain.	3, 4, 5	False certainty is absent. Model outcomes are confirmed by a post-hoc validation study. The case study focused on choosing a solution based on the state of the art data and models. Model application and available data involved uncertainties, but the uncertainties did not hamper choosing a solution.
Only risk assessment might be sufficient, e.g., when risks are low and solutions are not needed.	3	Risk assessments proved insufficient for the long term. The case study combined novel policy principles, risk- and sustainability metrics, and addressed a complex situation, which appeared to require SfSA to be solved and adopted.
Might not be applicable when there is no consensus on the problem, or whether there is a problem at all.	3, 4	Large pre-assessment available. Stakeholders were aware of and agreed on the large and urgent societal challenge, as well as on the meaning of a quantitative pre-assessment (phase-1 results).
Might result in investigating inappropriate solutions.	3	A sanity check is part of the procedure, both to check proposed solutions initially as well as via validation studies, monitoring and adaptive management.
May underestimate the complexity of cause-effect chains and thus the identification of management options	4	Too simple cause-effect chains (e.g. the application of too conservative safety margins) are identified and discarded (Posthuma et al., 2008). Agreement on systems-based cause-effect chains and the use of newly developed methods for quantification of local risks. Their clarification and acceptance in other fields supported their application in the new policies.

Acknowledging the (dis)advantages of Table 2.5, the final process design and the policy adoption of the results created by it, suggests that the application of the SfSA-principles has been necessary for solving the wicked environmental management problem.

In our experience, three aspects require specific attention in a successful SfSA. First, the pre-assessment is key. The pre-assessment defines the type of problem, according to the risk governance framework, but also requires a check on the basic assumptions, or 'first principles', of the present policy and regulation. For example: is sediment correctly handled in regulations as 'waste', or should it be considered a valuable resource? Moreover: is the policy (e.g. the Classification system) based on correct information (e.g. real risks, or on hazards only)? Existing regulations aimed at soil protection may have become a perceived or even legal blockage to choose a solution-focused innovative and more sustainable approach. That is why, during the pre-assessment, the problem is investigated from different disciplines, including lawmakers and lawyers. However, a clear distinction must be made between actual and perceived blockage. Perceived blockages are based on interpretations of the regulation rather than on the original intentions and underpinning of the regulation, for example interpreting a concept like Class-2 sediments as a 'risk fact' instead of a (conservative) decision tool to support daily practices. The phenomenon of interpretation of a 'concept' in terms of facts is known as 'reification' (Bradbury, 1989). Reification is a problem with substantial adverse consequences for practices when going unnoticed (Hyman, 2010). Thus, although stakeholder perspectives are an important element in SfSA, during the pre-assessment all scientific and regulatory positions should be explored according to their first principles rather than to historically evolved interpretations and grown implementation habits.

Secondly, adopting SfSA does not necessarily imply that existing methods are invalidated. For example, comparison of measured concentrations with generic soil, sediment and water quality criteria have been, and still are, a successful 'first defence line' approach against environmental deterioration. Whilst this quality criteria approach can remain to be a conservative first tier, applicable to- and successful for many cases to the aim of environmental protection, wicked cases can be solved with second- or higher tier approaches (Solomon et al., 2008), such as SfSA. As such, the concept of tiering can serve as a method to bridge the gap between problem analysing and problem solving principles: simple and conservative methods in the lower tier (to the end of protection, not invalidated), expanded with refined, system-specific and more precise methods in a higher tier (see Posthuma et al., 2008).

Finally, application of SfSA does not result in a single rule or solution but in an early sustainability optimization step, followed by adaptive management. For the sediment problem, this has occurred. There is evidence for ongoing optimization in some areas, including further stakeholder participation in decisions on sediment applications (Nieuwenhuis and Ellen, 2007). In such a process, regional bodies, such as water boards, can optimize area-specific sustainability metrics. Due to the SfSA approach, the focus in sediment management changed from sediment contamination to a system approach in which emission reduction can be combined with useful application of sediments in dikes and constructions and the realisation of 'nature-friendly' ditch shapes with all its benefits (e.g. lower sediment formation rate).

The needs for SfSA seem to grow. Global and regional environmental outlooks suggest an increasing incidence of wicked environmental problems, with trade-offs between water, energy, land use and food production metrics and conflicting social, economic, and environmental objectives (Sayer et al., 2013). Various proposals to deal with these problems exist. For example, Sayer et al. (2013) provide 10 principles to address wicked problems, many of which are shared with SfSA. Further, landscape approaches, instead of per chemical approaches, are proposed by Van der Horn and Meijer (2015) and Karlsson et al. (2011). And Sala et al. (2013) provide a meta-framework for sustainability assessments. SfSA builds forth on these existence frameworks, but specifically adds steps, principles and checkpoints related to the early position of the solution-focused paradigm in the assessment process.

2.5 Conclusions

Expanding on the idea of SfRA, we designed a versatile framework for sustainability assessment, SfSA, combining existing practices with the solution-focused paradigm. SfSA is innovative by the exploration of alternative management solutions upfront in the risk- or sustainability assessment process. Stepwise guidance to bring SfSA into practice has been elaborated. Based on a case study we conclude that SfSA can be supportive in successfully deriving and implementing sustainable solutions to wicked problems. It may help to break inertia, in cases where endeavours to solve wicked environmental problems with classical approaches fail. Based on our experience from the case study, the advantages of SfSA outweigh the disadvantages mentioned in the literature (Table 2.5). Further discussion and expansion of the SfSA principles and framework should be based on application in case studies relating to wicked problems. For SfRA, such case studies currently either lack, or do not cite SfRA literature sources. By creating a process in which stakeholders work together to find solutions, SfSA can help to build bridges between stakeholders and policy makers. SfSA needs not be applied always, given the strengths and successes of many easier applied existing approaches. SfSA can help, however, providing higher-tier support in more intricate multi-faceted problems.

2.6 Acknowledgments

This design of the SfSA framework was funded by RIVM's strategic research program, run under the auspices of the Director-General of RIVM and RIVM's Scientific Advisory Board, under project S/121001 and S/607020, and by the SOLUTIONS project. The SOLUTIONS project is supported by the European Union Seventh Framework Programme (FP7-ENV-2013-two-stage Collaborative project) under grant agreement no. 603437. The case study was funded by the Dutch Ministry of the Environment. We thank Sandra Boekhold, Irene van Kamp, Mark Huijbregts and the two anonymous reviewers of Environment International for their comments on draft versions of this manuscript.

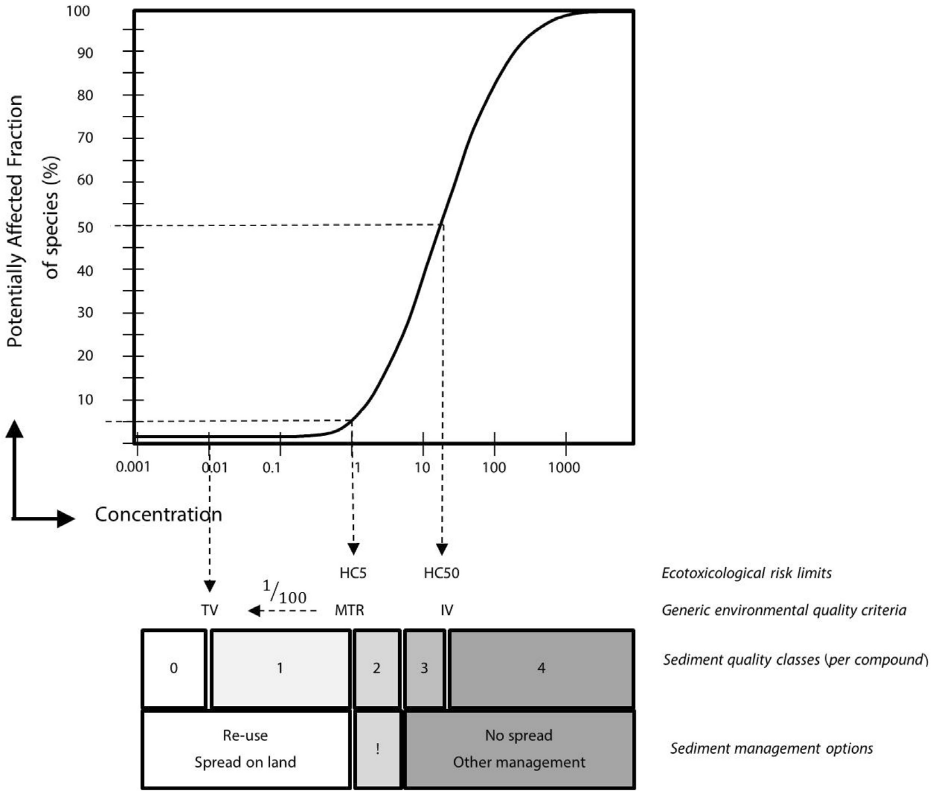


Fig. 2.7 Schematic representation of the contaminant-risk based principles underlying the management of sediments in the Netherlands since the 1990s (Ministry VROM, 1993). For explanation: see text. Environmental quality criteria for selected compounds are: TV=Target Value, MTR=Maximum Tolerable Risk, IV=Intervention Value, all expressed as chemical concentration. Science-based ecotoxicological risk limits are defined as HCx (Hazardous Concentration for x percent of the species of ecological processes (Posthuma and De Zwart, 2014). The "!" mark at the level of the sediment management options for class-2 represents the policy ambition to phase out spreading on land in 2003.

2.7 Annex: Background information regarding the case study

In the Netherlands, publication of the Clean Water Act of 1970 marked the implementation of policies to counteract water quality deterioration and its consequences, including sediment contamination. Given a choice for risk assessment as the basic principle for environmental policies (Ministry VROM, 1988), measures counteracting environmental deterioration were mandated, founded on science-based environmental quality criteria (maximum concentrations of chemicals) for soil, sediments and water (Ministry VROM, 1990; Van de Meent et al., 1990). For sediments, this resulted in the distinction of five sediment contamination classes (Burton, 2002; Ministry VROM, 1993), which related to various management perspectives (Fig. 2.7). The sediment contamination classes were separated by ecology-based quality criteria, which eventually were used for protection of soil quality. Water quality management policies further aimed to phase out the formation of contaminated sediments completely by 2025, with a formal aim to cease spreading Class-2 sediments on adjacent land in 2003 (Ministry V&W, 1998). Details on soil and sediment quality appraisal methods are presented in Boekhold (2008), Swartjes (2011) and Swartjes et al. (2012).

Contrary to expectations, contaminated sediment kept forming diffuse chemical emissions (Kramer et al., 1997; Kramer et al., 1998; Van Steenwijk and Mol, 1996), with expected adverse impacts on soil quality (Van Dijk et al., 1998; Van Dijk et al., 1999). This compromised the planned 2003 ban on spreading Class-2 sediments, and evolved into a regulatory problem related to continuous sediment formation. The situation evolved into a novel, wicked problem, from 1993 onwards, as more and more trade-offs of alternative sediment management options became apparent.

Triggered by concerns on soil quality, the studies of Kramer et al. and Van Dijk et al. in the late 1990s can be seen as early approaches towards a refined hazard identification for sediment deposition on land, being the first step of a classical risk assessment according to the U.S. NAS (1983) 'Red Book'. The approaches focused on the origins and fate of chemicals in both water and soil. Next, upon increasingly recognizing the wickedness of the problem given the different stakeholder perspectives, various management options were formulated in a series of historical studies (Table 2.6).

T**able 2.6** Historical studies and approaches for sediment management in the Netherlands, scored in hindsight for major SfSA criteria. SfSA criteria are: A = solution-focused, B = Clear definition on what to sustain, C = Participative approach, D = Clear definition of system boundaries, E = Innovative. For descriptions and study references: see text refs: (1)= (Ministry V&W et al., 1997), (2) = (Coopers & Lybrand, 1997), (3) = (Veul et al., 2000), (4) = (Hofstee and Leenaers, 2002), (5) = (OVAM, 2005), (6) = (Posthuma et al., 2006a) and (Ministry VROM et al., 2003). - = does not meet the criterion, + = meets the criterion, ++ = strongly meets the criterion. Cumulative score done by score summation via: + = 1; ++ = 2.

Period	Approach	SfSA criteria						
		A	B	C	D	E	Net Score	Reference
1997	Managing river sediments	+	-	-	-	+	2	(1)
1997	Managing municipal sediments	-	+	+	-	-	2	(2)
2000	Exploring on-site disposal options	+	++	++	++	++	9	(3)
2002	Active sediment management	++	+	++	-	-	5	(4)
2005	Sediment disposal practices in Flanders	++	++	++	++	+	9	(5)
2006	Novel Dutch sediment policy	+	++	++	++	++	9	(6)

The solution principles that have been explored, evaluated in hindsight, differ in the degree of linking up with the principles of SfSA (see scores in Table 2.6). For the evaluation, the principle 'clear definition of sustainability' was evaluated via 'clear definition of system boundaries' (criterion D) and 'clear definition of what to sustain' (criterion B). The SfSA principle 'iterative' applied to all studies and is not included in the table. Six studies were selected, which include at least one viewpoint on (the direction of) a solution. Application of SfSA principles differed across the studies, ranging from low to high fit (criteria sum-scores from 2 to 9). Scientific and policy attention was initially (1997) focused on the question how to prevent contamination of adjacent soils, lacking attention for the other sustainability domains (people and profit/prosperity respectively). Stakeholder participation was mainly limited to governmental stakeholders, who adopted the sediment classification system as starting point to define both the problem (expected transgression of soil quality criteria) as well as the solution (applying the classification system as a starting point for exploring novel solutions). Over the years, increasing insights in risk assessment techniques as well as in the complexity of the problem resulted in various creative solution options. Moreover, there was increasingly more room for stakeholder participation (criterion C) and innovation (criterion E, Table 2.6).

Over time, the approaches included more SfSA elements, without a consequent evolution regarding increased use of SfSA-elements. The relatively high score for the 2000 study (score 8) compared to the relatively low score of the 2002 study (score 5) relates to the type of study, being strategic explorative (listing possible future solutions) and pragmatic (listing an actual solution).

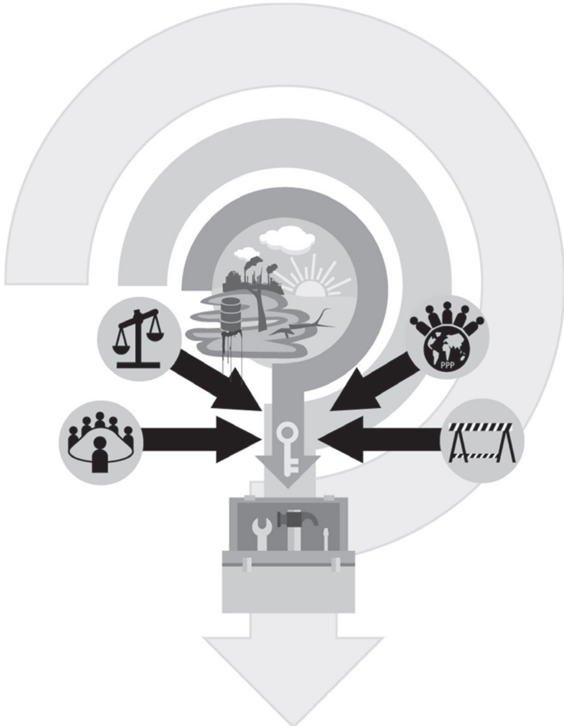
Pre-assessment is a key element of SfSA, and a need for emphasis on this step is substantiated in the history of the case study. That is, two technical aspects of the hazard definition could have been considered as basis for innovative solutions by detailed pre-assessment. One: Class-1 sediments were undefined regarding PAH concentrations in soils (Ministry VROM (1997), and this resulted in classifying vast amounts of Class-2 sediments at low PAH loads. In that way, hazards of sediments were implicitly over-estimated, without that being noticed. In communication, this type of error has become known as reification, a concept that is interpreted as a fact. Two: water and soil policies were founded on different principles, i.e., a system-approach for water (related to hydrological systems), and a generic, protective approach for soil, respectively. In the context of evolving scientific insights and practices, it was recognized that an improved pre-assessment could form the basis for new policies. From a scientific perspective, it was recognized that bioavailability, mixture toxicity and compound degradation in aerobic soils should be taken into account (Peijnenburg et al. (2005)), implying again that risks of sediments classified as Class-2 due to e.g. PAHs would likely be lower than anticipated. Further, validated approaches to take these issues into account became available (De Zwart and Posthuma, 2005; De Zwart et al., 2008). Regarding regulations, it was recognized that environmental quality management could be founded on an integrated water-sediment-soil systems assessment approach, with a quantitative assessment of risk and sustainability metrics (Posthuma et al., 2006a). The quantitative approach could be envisaged as a second-tier approach, following a comparison with soil quality criteria in the first tier, instead of replacing them. The latter recognition eventually resulted in adoption of novel policies, supported by a broad selection of stakeholders (Osté et al., 2008; Wezenbeek et al., 2007).

The study underlying the novel policy and its acceptance and implementation (study 6 in the Table) scores positively on many SfSA elements. The pre-assessment consisted of a comprehensive study on the national and regional sediment loads and sediment formation masses (AKWA, 2001), an exploration of the impact of the scientific innovations (quantitative risk assessment of human and ecological risks), and a collation of stakeholder views on requisites for solutions. The system boundaries were evaluated and defined, and protection goals were transparently defined. The procedure was participative, and considered multiple viewpoints (environmental aspects, human interest and economical boundary conditions). Consideration was given to views from stakeholders composed of national and regional authorities in land and water management, food producers, food sellers and food traders and scientists with expertise in chemical risk assessment and regulation, soil, sediment and water risk assessments and management. Communication was arranged on a regular basis, to link the various SfSA steps and to inform stakeholders and discuss progress. Due to elapsed time since problem recognition, the process was characterized by urgent needs to solve the problem. The final solution was evaluated to be societally sound and financially realistic, without compromising soil quality (given field validation results) or risk increase.

3 An identification key for selecting methods for sustainability assessments

Authors: Michiel C. Zijp, Reinout Heijungs, Ester Van der Voet, Dik Van de Meent, Mark A.J. Huijbregts, Anne Hollander, Leo Posthuma

Published in: Sustainability (2015) 7: 2490-2521



Abstract

Sustainability assessments can play an important role in decision-making. This role starts with selecting appropriate methods for a given situation. We observed that scientists, consultants, and decision-makers often do not systematically perform a problem analyses that guides the choice of the method, partly related to a lack of systematic, though sufficiently versatile approaches to do so. Therefore, we developed and propose a new step towards method selection on the basis of question articulation: the Sustainability Assessment Identification Key. The identification key was designed to lead its user through all important choices needed for comprehensive question articulation. Subsequently, methods that fit the resulting specific questions are suggested by the key. The key consists of five domains, of which three determine method selection and two the design or use of the method. Each domain consists of four or more criteria that need specification. For example in the domain "system boundaries", amongst others, the spatial and temporal scales are specified. The key was tested (retrospectively) on a set of thirty case studies. Using the key appeared to contribute to improved: (i) transparency in the link between the question and method selection; (ii) consistency between questions asked and answers provided; and (iii) internal consistency in methodological design. There is latitude to develop the current initial key further, not only for selecting methods pertinent to a problem definition, but also as a principle for associated opportunities such as stakeholder identification.

3.1 Introduction

The recognition that mankind puts major pressures on the earth systems has resulted in publications reporting on the results of sustainability assessments, e.g. (Hoekstra and Wiedmann, 2014). Not only environmental crises, but also social inequities at local to global scales trigger a societal drive to position sustainable development as a decision-making strategy (Waas et al., 2014). There is a call for sustainability assessments from the local scale, such as sustainable development of cities and neighborhoods (ICLEI, 2014), to the global scale, for example the United Nations sustainability goals (UN, 2012); and from the product level, e.g., eco labels(EC, 2013), to the sector level (EC, 2014a). This demand resulted in a plethora of methods that claim to provide answers to sustainability questions. In fact, we have entered an era in which there is an abundance of methods for sustainability analyses. Some of these methods are complimentary, but there are also many competing methods. In the meantime, sustainability science is a swiftly developing discipline. There is an ongoing debate on what sustainability or sustainable development is, and what sustainability assessment should encompass, whilst there is a need to bridge widely diverging disciplines, each with their own definitions and approaches. In this important, complex and swiftly developing field, the selection of appropriate methods for answering a particular sustainability question can be challenging.

Whilst scientific approaches ideally are "fit for use" and robust, the selection of sustainability assessment methods is thought to be frequently led by the expertise of the analyst and available capacity (Gasparatos and Scolobig, 2012; Sala et al., 2013). The choice of method(s), however, requires choices on scope, assumptions, values and precision. One question regarding sustainable development can therefore yield a manifold of answers, depending on which assessment method is selected to answer the question (Browne et al., 2012; Özdemir et al., 2011). If the underlying choices in method selection are not considered explicitly, there is a chance of a mismatch between the results and the context in which the question was asked. Also, the answer becomes more dependent on the analyst to whom the question is directed, than on a transparent specification of the question and methods logically linked to such a question. Fundamental improvements in the analysis of sustainability questions seem warranted, given these observations.

Looking at optimal improvements, literature shows that there are attempts to organize available sustainability assessment methods. These primarily should support the selection of appropriate method(s) to address the problem. First, lists have been drawn of sustainability methods, including reviews per method on objectives, strengths, weaknesses, et cetera (De Ridder, 2005; Singh et al., 2012). Although useful to gain insight in all current optional sustainability assessments, such lists and descriptions provide limited guidance to selection of a method. They are not designed to compare methods, but to provide information per method. Because of the amount of methods and their details, the lists contain a substantial amount of information. The second type of organization is that methods are described or scored based on a selection of criteria that are found to be important, such as object and spatial focus (e.g., Finnveden and Moberg (2005); Wrisberg et al. (2000)). Different sets of criteria are used for this purpose. An overview of these criteria, based on 26 studies, can be found in Table 3.1. The list of potential criteria based on which one can distinguish between methods is long and the possible directions per criteria vary. This approach provides a structured way to organize all the information gathered per method. Still, it is a lot of information to process, since a comparison of all criteria for all methods is required in order to select the most suitable method for a question. As variant of the importance scoring, and to reduce the amount of information per method, some authors appoint a selection of their list of criteria as key-criteria. The basis for this selection, however, is often not found in the manuscripts or supporting information, e.g., (Finnveden and Moberg, 2005; Hacking and Guthrie, 2008; Sala et al., 2012b). An exception is found in Udo de Haes et al. (2006) who, for the purpose of comparing Life Cycle Assessment (LCA) and Ecological Risk Assessment (ERA), give an argumentation per criterion why it is either a fundamental or secondary criterion. The third approach to organize methods is to categorize or frame them, based on a selection of features of the methods. Some much cited attempts of categorization can be found in literature (e.g. Ness et al. (2007). Although categorization might be valuable to gain a quick insight in the type of methods available, the approach provides limited support for method selection. Firstly, because the loss of information: the variability between methods is too large to be captured in three or four features. Secondly, because methods themselves are flexible and can often be applied in various modes for different purposes. This struggle with complexity and flexibility can for example be seen in the frequently cited categorization of Ness et al. (2007). Their framework provides a categorization based on, amongst others, the object of the sustainability analysis, but it has only room for the objects "spatial unit" and "products". Further, various methods positioned in the framework would also fit on other places in the framework. For example, LCA could also fit under "integrated indicators" and all the methods under the integrated "prospective" categorization fit under "retrospective" as well. Other attempts for categorization frameworks (De Ridder et al., 2007; Gasparatos and Scolobig, 2012; Hacking and Guthrie, 2008; Sala et al., 2013) show the same type of struggles with existing variability and complexity of sustainability assessments. They are useful for a quick overview on the type of assessment methods that are available, but provide insufficient detail for method selection.

Scrutiny of the aforementioned approaches shows that they are often supply-driven. Methods used in sustainability assessments are organized based on the articulation of the available methods. Ideally, however, the selection of a method for answering a specific sustainability question would be performed based on a specific analysis of the assessment problem, and a subsequent specific articulation of the research question (demand-driven), which would result in explicit choices for assessment method(s). Some studies do experiment with an analysis of the assessment demand as organizing principle. For example, Wrisberg N. et al. (2000) state that every demand has its specific object of analysis, spatial and temporal dimensions, required level of detail and required level of integration, which should be leading in choosing a method or a combination of methods. Furthermore, they distinguish five type of contexts (strategic, capital investment, design and development, communication and marketing and operational) and eight context specifications with each their specific profile of demands for method selection. Another example is from De Ridder et al. (2007), who analyzed which phase in integrated assessment frameworks, like Strategic Environmental Assessment, asks for which type of method. When linking demand with supply, both De Ridder et al. and Wrisberg et al. show that the context (e.g., the phase in the management cycle) does not seem to be leading in method selection, but rather in method design. The context specifies the role and design of a selected method, such as the thoroughness of the analysis or the way results are presented, and not the method selection itself (Van Passel and Meul, 2012).

Many examples exist in which method selection is based on explicit qualitative comparison (description of strengths and weaknesses) or characterization. These approaches are, however, case specific and do not cover the extent of relevant criteria and methods to choose from, e.g., (Binder et al., 2010; Browne et al., 2012; Carof et al., 2013; Florin et al., 2012; Özdemir et al., 2011).

Recent literature on novel ways to support method selection in the field of sustainability assessment is scarce. A recent innovation is described by Gasparatos and Scolobig (2012), who provide four different proposals to base method selection on (Gasparatos and Scolobig, 2012), namely based on:

1. the perspective of the assessment (e.g., biophysical limits or human wellbeing);
2. desired features of the assessment (e.g., spatial or temporal focus);
3. the acceptability criterion of Pope et al. (2004) (Pope et al., 2004) (e.g., is the goal of the assessment to reduce impacts or to reach explicitly defined sustainability goals?);
4. values of the stakeholders (e.g., focus on general human well-being, personal well-being, or ecosystem well-being).

Given apparent limitations of current approaches in sustainability assessment method selection, we propose a next step. A next step in this field would be a framework for method selection that takes into account all four proposals of Gasparatos and Scolobig (2012), and possibly more, but that is also capable of capturing the dynamics and complexity of both the question (demand) and method (supply) side of sustainability assessments. We propose to collate all aforementioned approaches and options into an identification key, inspired by e.g., Flora's for plant determination. We expected that and evaluated whether this could be the step forward into systematic and transparent method selection. We expect that the provision of a Sustainability Assessment Identification Key (SA-IK) is of general support to:

I. identify a method based on explicit choices and all methods available and not necessarily the well-known method by the analyst;
II. guide method selection from demand perspective (articulation of the question) rather than supply perspective;
III. report the results of the assessment referring to the explicit choices made with question articulation, making results easier to understand, interpret and compare with other assessments.
IV. make method selection transparent and reproducible.

Before these features can be substantiated, such an identification key had to be designed and tested. In this context, the aims of this study are:

I. to confront the available assessment methods with the sustainability questions posed by society such as to propose a new organizing framework for selection of sustainability assessment methods: the sustainability assessment identification key;
II. to present the design of the sustainability assessment identification key;
III. to show how the sustainability assessment identification key (SA-IK) works.

It should be noted that the emphasis of this study, the SA-IK design and its application, mainly considers environmental assessment methods, focusing thus on the "planet" aspect of sustainable development. Further, we acknowledge that sustainability assessments may reflect subjective views of researchers on the definition of sustainability, and that sustainability assessments can differ due to that. Independent of such views, the SA-IK was designed to support clear and transparent support in choosing and applying methods. When the SA-IK is potentially used by various types of users, those users are expected to specify how sustainability is defined by them for each case specifically, transparently and explicitly.

3.2 Methods: Development of a Sustainability Assessment Identification Key

3.2.1 Terminology used in this article

The scientific literature in which sustainability assessment methods are being described and characterized shows a plethora of terms such as instrument, tool, method, methodology and procedure, which are more or less interchangeable, but having different meanings in different articles. In practice a certain degree of hierarchy can be discovered between them (Sala et al., 2012b), see Fig. 3.1. The definitions used in this paper are described below.

A method is defined as a collection of consecutive and complementary sub-methods with which a specific question can be answered. Examples of methods are Life Cycle Analysis (LCA), Ecological Footprint (EF), EMergy analysis (EM), but also indexes like the Dow Jones Sustainability Index. Consequently, sub-methods are the consecutive and/or complementary analytical steps a method consists of. For example methods with which indicators can be quantified or with which results can be aggregated. The distinction of methods and sub-methods is important, because in practice methods do not necessarily have to be applied as a whole. For example, LCA contains a sub method with which emissions and the use of resources can be translated into impacts: the Life Cycle Impact Assessment (LCIA). This sub-method can (and is e.g., Schneider et al. (2015)) also be used to determine the impact of activities without taking into account the life cycle of that activity, thus without taking the LCA method into consideration as a whole.

Procedure
Stepwise approach consisting of different phases of a process which provides the link with the societal and decision context (e.g. Strategic Environmental Assessment, Integrated Sustainability Assessment).

Framework
The rationale and the structure for the integration of concepts, methodologies and methods (e.g. Sustainability Science conceptual framework, Life Cycle Sustainability Assessment, …).

Methods
A collection of individual consecutive and/or complementary sub-methods. Methods can include 1) indicator selection, 2) quantification or qualification and 3) aggregation, but not necessarily all three aspects (e.g. Life Cycle Assessment, Risk Assessment, Dow Jones Index, …).

Sub-methods
A method to agree on, estimate, calculate or aggregate results (e.g. delphi method, fate model, weighting).

- **Tool** is a brand that executes a method (software, application, …).
- **Indicator** is a parameter, or a value derived from parameters, which points to, provides information about, or describes the state of a phenomenon, with a significance extending beyond that directly associated with its value (OECD, 2003). The parameter could be quantitative or semi-quantitative or qualitative derived from a method or sub-method.

Fig. 3.1 Terminology and their hierarchical relation adopted in this paper. Adjusted from Sala et al. (2012b).

Methods are the operationalized parts of a higher-level entity which we refer to as the framework in which the sustainability assessment is undertaken. A framework is the way sustainability is conceptualized. For example: Which pillars (that is: people, planet, prosperity) are thought to be important? How do they relate to each other? And: How are interactions between impacts at different spatial and temporal scales envisioned?

The framework itself is again part of wider context, which we refer to as procedures. Procedures consist of subsequent phases in a process of making decisions and in policy, given the existence of a societal problem. Methods can have varying roles in different phases of a procedure (De Ridder et al., 2007).

While this study focuses on the selection of methods, also the other levels (framework and procedure) play an important role. Namely, to select and especially to design a method tailored to the problem definition, the place in the procedure and the framework envisioned by the user must be made explicit and thus known. In other words, choices that are made in sustainability assessments regarding method(s) choice are co-influenced by the contexts of the framework and procedure within which problems can be addressed.

3.2.2 Review on the derivation of an Identification Key in general

Identification keys exist in many fields, e.g., biology, medicines, social sciences and information architecture. These disciplines have in common that they intend to classifying large amounts of objects, at representing a complex system of characteristics, at easy adjustment when new insights or findings are added to the collation of objects, and that are—finally—of use to practitioners. In biology and other fields the result of an identification key is often referred to as a taxonomy of the objects under study. Although relevant for various disciplines, the available scientific literature on taxonomy development is limited (Nickerson et al., 2013). Three types of processes towards a classification, or taxonomy, have been distinguished. First, there is the conceptual approach in which a classification system is drafted based on a conceptual model or idea of the field of interest in a deductive manner and subsequently improved or tested based on e.g., empirical data (Bailey, 1994). The second approach starts with empirical data and builds a classification system based on statistical methods (ibid). Thirdly, a classification system can be based on interaction with users of the information that needs to be classified, which is also named a "folksonomy" (Peters, 2009). Expertise and experience on taxonomy in the fields of biology and knowledge engineering shows that there is not something like a perfect or ideal taxonomy for a field of interest. Complex information can be organized and unlatched in many ways. For example, in ecology, organisms are being characterized based on either physical characteristics (phenetics) or on evolutionary relationships (cladistics), resulting in significantly different groupings of organisms, but both function well as classification systems. Another lesson to be learned from taxonomic experiences in other fields is that designing a comprehensive, widely supported identification key can take years (Dijkers et al., 2014; Kendal and Creen, 2007). Deriving a meaningful identification key is an iterative process (Fig. 3.2). This will also hold for sustainability assessment, where basic principles of the field are still much debated (Sala et al., 2012a; Sala et al., 2012b; Waas et al., 2014). Taking that in mind, a first design of an identification key, like presented in this paper, should be seen as an essential, but small step towards a complete and widely applicable and applied key.

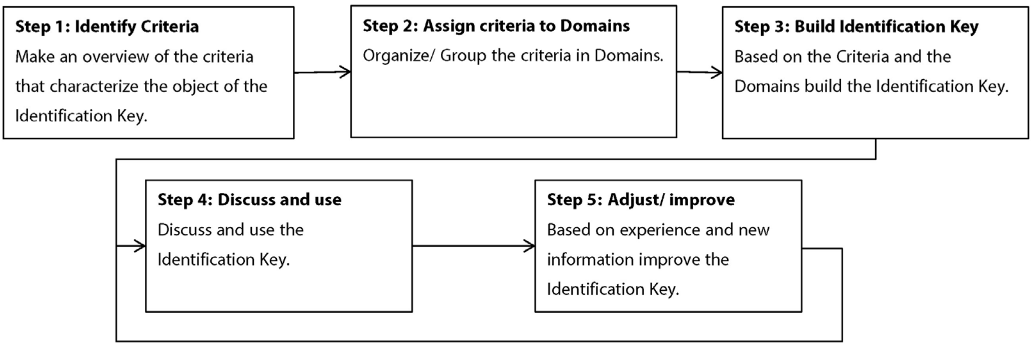


Fig. 3.2. The identification key development method based on Nickerson et al. 2013.

Where taxonomy is the art of organizing and unlatching available information, the identification key we intended to design should also deal with information that is not yet available. Namely, for some (or perhaps many) sustainability questions, a method to derive the assessment outcomes does not (yet) exist. Because potentially a large group of sustainability assessment methods are not known yet, a conceptual approach is followed to build the sustainability assessment identification key and not a statistical approach. Nickerson et al. (2013) provide a taxonomy development method, which we adjusted (Fig. 3.2) and used for the development of the Sustainability Assessment Identification Key (SA-IK).

3.2.3 Step 1: Identify criteria

As already mentioned, Table 3.1 presents an overview of the criteria used in 27 literature sources that focus on method selection, both from the supply perspective and the attempts to include the demand perspective. The column with references in Table 3.1 shows the difference in criteria sets that are thought to be important by the different sources. For the purpose of presentation criteria that consider similar concepts were grouped together under a single criterion and assigned a description, a step improving clarity, but removing nuance. For example, the criteria "values/view on sustainability" in Table 3.1 includes the choice of weak versus strong sustainability (see Table 3.1), the choice of world view (e.g., anthropocentric or ecocentric) and risk perception (e.g., Cultural Theory). In the identification key, these different aspects of values can be dealt with separately.

Table 3.1 List of criteria derived from literature sources and classified in different domains, with a short description and references per criteria.

Criteria	Explanation	References
Domain: System boundaries/Inventory		
Object	*What is the object of the assessment?* *Is it a physical object (product, chemical, process), or an organization, a region, a policy measure, an activity, etc…*	(Finnveden and Moberg, 2005; Jeswani et al., 2010; Ness et al., 2007; Singh et al., 2012; Udo de Haes et al., 2006; Wrisberg N. et al., 2000)
Spatial focus	What is the spatial focus of the activity? *Is the activity assessed on micro or macro scale, and if on macro on local, regional or global scale?*	(Jeswani et al., 2010; Ness et al., 2007; Parris and Kates, 2003; Sala et al., 2013; Singh et al., 2012; Udo de Haes et al., 2006; Wrisberg N. et al., 2000)
Temporal focus	What is the temporal focus of the assessment? *Is the activity assessed retrospective, prospective or does a snapshot suffice?*	(Gasparatos and Scolobig, 2012; Hacking and Guthrie, 2008; Jeswani et al., 2010; Ness et al., 2007; Pope et al., 2004; Sala et al., 2013)
Life cycle thinking	Which parts of the life cycle or supply chain are included in the assessment? *Only one phase, the whole life cycle, or something in between?*	(Mayer, 2008; Sala et al., 2013; Sala et al., 2012b; Udo de Haes et al., 2006)
Domain: Impact Assessment/Theme selection		
What is to be sustained	*What is to be sustained?* Are these environmental, social, economic and/or institutional endpoints?	(Böhringer and Jochem, 2007; Gasparatos et al., 2008; Hacking and Guthrie, 2008; Jeswani et al., 2010; Ness et al., 2007; Parris and Kates, 2003; Sala et al., 2013)
Theme and indicator selection	*Which themes are selected?* *Is the method transparent in the selection and use of indicators? What place on the cause effect chain do the indicators have? etc.*	(Böhringer and Jochem, 2007; Finnveden and Moberg, 2005; Joumard, 2011; Singh et al., 2012; Udo de Haes et al., 2006)
Domain: Aggregation/Interpretation		
Spatial focus of impact	*What is the spatial scale of the impacts that should be taken into account?* Does the assessment include intra-generational impacts? Or in other words: does the assessment aim at internal or external sustainability. Impacts at what scale are taken into account? Are they site-specific/dependent or independent?	(Hacking and Guthrie, 2008; Joumard, 2011; Mayer, 2008; Udo de Haes et al., 2006; Wrisberg N. et al., 2000)
Temporal focus of the impact	*What is the temporal scale of the impacts that should be taken into account?* Does the assessment include inter-generational impacts? What time-frame should be included for the impacts?	(Bond and Morrison-Saunders, 2011; Gasparatos et al., 2008; Gasparatos and Scolobig, 2012; Hacking and Guthrie, 2008; Parris and Kates, 2003)

Criteria	Explanation	References
Domain: Aggregation/Interpatation		
Sustainability target	*Is a sustainability target necessary?* If the goal is to compare alternatives, to perform a hotspot analysis or to improve an object, a sustainability target is not essential. If the goal of the analysis is to determine the sustainability of an object, a target is required. This is also referred to as direction to target (no target needed) or distant from target (target needed); and assessment impact-led (least impact, no target needed), objective-led (best positive contribution, no target needed) or assessment for sustainability (like the other two, but in relation to a specific sustainability target)?	(Gasparatos and Scolobig, 2012; Hacking and Guthrie, 2008; Mayer, 2008; Pope et al., 2004; Sala et al., 2013)
Values/View on sustainability	*What view on sustainability should be leading in the assessment?* Is sustainability understood as weak, strong or partly substitutional? In short: weak means that various capitals are interchangeable. Strong means that each capital should be preserved independently. Partly substitutional means weak until a critical level is reached, e.g., Critical Natural Capital (CNC) or planetary boundary. Also one's world view (personal believes or risk perception) can influence the assessment.	(Bond and Morrison-Saunders, 2011; De Schryver et al., 2011; Dietz and Neumayer, 2007; Gasparatos et al., 2008; Gasparatos and Scolobig, 2012; Özdemir et al., 2011; Pope et al., 2004; Robèrt et al., 2002; Sala et al., 2013; Singh et al., 2012; Svarstad et al., 2008; Zoeteman, 2001)
View on integration of pillars	*How should aggregation of information from different disciplines take place in the assessment?* In a multi (separate), inter (connected) or trans (combined/holistic) disciplinary way?	(Bond and Morrison-Saunders, 2011; Gasparatos et al., 2008; Gasparatos and Scolobig, 2012; Hacking and Guthrie, 2008; Ness et al., 2007; Sala et al., 2013; Sala et al., 2012b)
Normalisation/ weighting/ aggregation method	*Which aggregation level is preferred and which methods are used?* Both normalisation (make data comparable), weighting (specify interrelationships) and aggregation (get functional relationships) need careful consideration.	(Böhringer and Jochem, 2007; Gasparatos et al., 2008; Mayer, 2008; Özdemir et al., 2011; Parris and Kates, 2003; Singh et al., 2012; Udo de Haes et al., 2006)
Domain: Method Design		
View on stakeholder involvement	*Who should be involved in the assessment in which way?* Also referred to as legitimacy, in relation to indices or composite indicators.	(Bond and Morrison-Saunders, 2011; De Ridder et al., 2007; Gasparatos et al., 2008; Gasparatos and Scolobig, 2012; Parris and Kates, 2003; Robèrt et al., 2002; Sala et al., 2013; Thabrew et al., 2009)
Context of the assessment	*How and by whom are the results used?* In which (phase of a) procedure are the results of the assessment used? Is the goal of the measure: decision-making and management, advocacy, participation and consensus building or research and analysis? Or is it a strategic, capital investment, design and development, communication and marketing or operational question?	(De Ridder et al., 2007; Finnveden and Moberg, 2005; Jeswani et al., 2010; Parris and Kates, 2003; Sala et al., 2013; Wrisberg N. et al., 2000)
Uncertainties	*How are uncertainties to be handled?* Salience, credibility and variability? Should an uncertainty, sensitivity and/or perturbation analysis be included?	Gasparatos et al., 2008; Gasparatos and Scolobig, 2012; Hacking and Guthrie, 2008; Parris and Kates, 2003; Pintér et al., 2012; Sala et al., 2013; Wrisberg N. et al., 2000)

Criteria	Explanation	References
Domain: Organisational restrictions		
Formal requirements	*Should the method be formally recognized?* ISO, EC, etc.	(Parris and Kates, 2003; Wrisberg N. et al., 2000)
Expertise requirements and availability	*Is there capacity for hiring expertise?* Expertise requirements and availability	(Gasparatos et al., 2008; Singh et al., 2012; Wrisberg N. et al., 2000)
Software requirements and availability	*Is there capacity for acquiring software?* Software requirements and availability	(Singh et al., 2012; Wrisberg N. et al., 2000)
Data requirements and availability	*Is there capacity for gathering data?* Data requirements and availability	(Mayer, 2008; Olsen et al., 2001; Parris and Kates, 2003; Wrisberg N. et al., 2000)

3.2.4 Step 2: Assign the criteria to domains

For the design of the SA-IK we distinguished three domains that determine the method selected for a specific sustainability question, and two that determine the further design and use of the method. The first domain deals with the system boundaries of the activity under consideration or, in other words, the specification of the inventory (in LCA terms) or system quantification (in Material Flow Analysis terms) or Drivers and Pressures (in DPSIR terms; DPSIR is further explained in section 2.6). It contains question articulation on, amongst others, the type of object, the spatial scale and other criteria that determine the system boundaries of the assessment. This first domain is further referred to as "System boundaries/Inventory". The second domain: "Impact assessment/Theme selection"; articulates the type and scope of the impacts. For example, which themes or issues are thought to be important? Is the focus on environmental issues, or also on social, economic and institutional issues? And: is the focus on impacts on one specific location or impacts worldwide? The third domain articulates the need and specification for aggregation of the results and is named: "Aggregation/Interpretation". The idea to distinguish these three domains for method selection is based on the observation that sustainability assessments methods in general consist of these three elements. For example, in LCA the activity and its resulting emissions and resource uses are quantified based on a set of boundary conditions in the inventory phase, which is followed by the impact assessment phase in which the emissions and resources used are expressed in themes (impact categories) that are thought to be important. Finally, choices are made on if and how the results per theme need to be aggregated, e.g., by weighted summation. Another example of this triptych (1. system boundaries/inventory, 2. impact assessment/theme selection, 3. aggregation/interpretation) is (societal) cost-benefit analysis (CBA): CBA starts with defining the policy alternatives in a given situation (1. system boundaries), followed by translating these activities and its results in costs and benefits for different stakeholders (2. impact assessment). Then, choices are made in how the results are presented (aggregated or not) and weighted (3. aggregation and interpretation).

The domains are chosen such that the Criteria found in literature (step 1, paragraph 2.3) occur in just one of the domains. For example, the role of personal "values" for method selection (is ones view on sustainability strong, weak or partially compensatory) are influential for the choice of aggregation method, but less on the choice of system boundaries and theme selection.

Some of the Criteria derived from literature (Step 1, paragraph 2.3) do not determine the selection of methods, but rather steer the design or use of a method, e.g., the inclusion of uncertainty analyses or stakeholder involvement. For these Criteria we added the domain "method design". Other Criteria are organizational restrictions provided by other influences then the question or the context of the use of its results. We refer to these Criteria as organisational restrictions. Table 3.1 groups the Criteria per Domain. Further details can be found in the Supporting Information (http://www.mdpi.com/2071-1050/7/3/2490#supplementary, last visited 09-03-2017).

3.2.5 Step 3: Build the Identification Key

Based on the Domains and Criteria specified above, a first design of the identification key for sustainability assessment was derived. Conceptually, the drafted SA-IK consists of a key that focuses first on the three domains System boundaries/Inventory, Impact assessment/Theme selection and Aggregation/ Interpretation. When the Criteria in these Domains are addressed, the SA-IK delivers as main outcomes an articulated question and suggestions for method(s) selection per domain. That is, methods that fit the chosen boundary conditions, methods that fit the chosen themes and methods that fit the specifications regarding aggregation. The outcomes of these three keys are then collated and analyzed, yielding one selected method, a (new) combination of (sub) methods or the conclusion that a method for the specified question is not (yet) available in the SA-IK. The selected method(s) can be further designed with the help of the subsequent Method design key, for example to add a sensitivity analysis. Moreover, the Criteria labeled "organizational restrictions" (pragmatic constraints, like time and data availability), can influence the methods choice and design, but are left out of the scope of this first design.

As an example, Fig. 3.3 visualizes the identification key for the part focusing on System boundaries/ Inventory. The first question in the identification key is: what is the Object of the analysis? This can be products, geographic units, companies, et cetera. The follow-up question depends on the answer given. For example, the first question for the Object "products" could be: "Does the situation concern (a) single product(s) or (a) product group(s)?" whereas the first question for the Object "geographic unit" could be: "what is the spatial focus of the activity?" and so on.

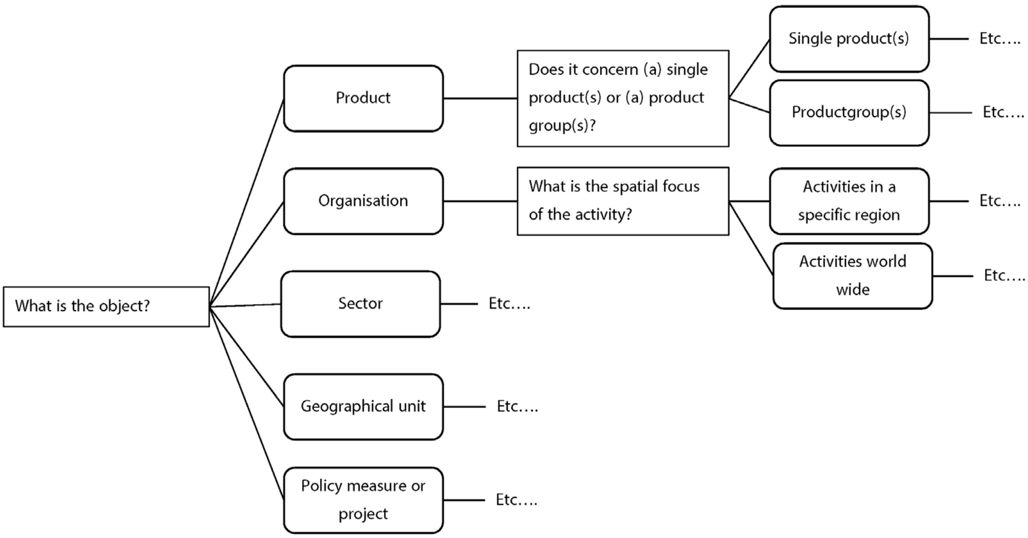


Fig. 3.3 View of the first part of the identification key for the Domain System boundaries/Inventory as an example.

3.2.6 Note on theme selection

The previous step results in a systematic evaluation of choices and evaluations relevant for selection of a method for sustainability assessment. Next to literature on method selection, on which our analysis and IK primarily focuses, an even larger amount of literature exists on theme and indicator selection. Themes are the issues that could be considered to take into account when looking at the sustainability of an object. Indicators are parameters that provide information about (or describes the state of) these themes (OECD, 2003); see also Fig. 3.1. For example, concentrations of nutrients in fresh water (the indicators) can be an indicator for ecological damage due to eutrophication (the theme). Indicators can be chosen on different places at the cause-effect chain. The DPSIR framework is often used to describe the relation between an indicator and the Driver (the activity), the Pressure (e.g., the resulting emissions), the State (e.g., the concentrations in fresh water), the Impact (e.g., on biodiversity) and the Response (e.g., policy measure or monitoring) (Niemeijer and De Groot, 2008b), mostly in environmental assessments, but also broader (Blok et al., 2013; Svarstad et al., 2008).

Niemeijer and de Groot (2008a) provide an overview of criteria found in literature for indicator selection, which will not be repeated here, but which could be useful for a themes and indicators identification key, to be used following the SA-IK design. Although theme selection and method selection are closely related and always needed both, they are often treated as separate entities. This results in two approaches for designing a sustainability assessment. The first approach is that one or more methods are chosen, followed by theme selection. The second approach is that first themes are chosen, followed by a selection of methods with which these indicator representing the themes can be quantified. The drawback of the first approach is that existing methods are often limited in that they represent a current set of specific themes (and not: all potentially relevant themes), and thus the choice of method narrows down the options of considering potentially relevant themes (when uncritically applied). The second approach starts with theme selection and gathers methods able to quantify the selected themes. Though this suggests a problem-driven choice of theme, and thus relevance of the final results, this working order also has potential drawbacks. Firstly, there is the probability that the results for the various themes are less comparable (e.g., life cycle based and not life cycle based; site dependent versus site independent impacts). Secondly, one runs the risk of applying methods in a wrong matter (e.g., linear extrapolation of LCA results from micro to macro level) to be able to compare the results of the different themes. Thirdly, there is the risk of double-counting when aggregating the results (e.g., both the "Pressure" and the "Impact" as indicator). Scientific literature reports many examples of both alternative approaches: method selection is leading versus theme selection is leading.

3.3 Results: how the Identification Key works and the problems it solves

3.3.1 Example of Sustainability Assessment Identification Key application

Next to designing SA-IK, we aimed to test and improve it iteratively. The key we present has been subject to this iteration, and the final SA-IK (step 5 in Fig. 3.2) was used to illustrate its use and usefulness, according to our third study aim.

First, we explored how systematic use of the SA-IK in the phase of sustainability assessment question articulation can result in method selection, based on the first design of the sustainability assessment identification key (SA-IK). An example of the results of this exploration is provided in Table 3.2. The table show three ways to articulate a realistic societal problem. Thus, three teams start with the same general question, but refined question articulation guided by SA-IK leads them to three different specific questions and thus three different methods selected. The focus of the example is not on why certain choices are made, but on what the consequences of the choices are for the method selected and thus the type of answer provided. The SA-IK provides the elements that require a choice and a general list of possible answers per choice. For example, the question in Table 3.2 is: "How sustainable is our food pattern?" The first choice provided by the SA-IK is: What is the object? This could be products consumed, but also lifestyles, the foodsector, foodpolicy, or a certain geographic unit. This choice, and its possible directions, are discussed and decided on by the actors responsible for the question articulation. The following choice depends on the answer to the previous choice. For example: the object "product" should be further specified in "single product(s) or product groups"?; a question that is not relevant for the object "geographic unit". In that way, the SA-IK guides its user through explicit choices regarding all criteria in the System boundaries domain and then provides a list of methods that fit the choices made. The Impact assessment/Theme selection and Aggregation/Interpretation domains are specified similarly followed by comparison and selection of the method(s) that the SA-IK found suitable for answering the articulated question.

Table 3.2 Three examples on how the Sustainability Assessment Identification Key (SA-IK) applied to an apparently singular question leads to different method choices, given specifications identified by explicitly addressing SA-IK Domains and Criteria. The examples start with the same question, but they follow different contextual–and therefore assessment articulation–pathways, leading to different questions to be answered by different methods.

Example 1		**Example 2**		**Example 3**	
Question	How sustainable is our food pattern?	**Question**	How sustainable is our food pattern?	**Question**	How sustainable is our food pattern?
		Sub Identification Key on System boundaries/Inventory			
What is the object?	Products	**What is the object?**	Products	**What is the object?**	Geographical unit (river catchment)
Single product(s) or product group(s)?	Single products	**Single product(s) or product group(s)?**	Product groups		
Should the product(s) life cycles be included?	Yes	**Should a chain analysis be included**	Yes	**Should a chain analysis be included?**	Yes
Which part of the life cycle?	Cradle to grave	**Which part of the chain?**	Upstream	**Which part of the chain**	Upstream
What is the spatial focus of the activity	Local	**What is the spatial focus of the activity**	Regional	**What is the spatial focus of the activity**	Continental
What is the temporal focus of the activity?	Snapshot	**What is the temporal focus of the activity?**	Snapshot	**What is the temporal focus of the activity?**	Prospective
Results sub IK System boundaries/ Inventory	Life Cycle Inventory		Input Output Analysis, Material Flow Analysis, Substance Flow analysis, …		Input Output type of analysis in combination with scenario building
		Sub Identification Key on Impact assessment/Theme selection			
What is to be sustained?	Environment	**What is to be sustained?**	Resources	**What is to be sustained?**	Biodiversity
Which location on the cause effect chain is required?	Impact at endpoint	**Which location on the cause effect chain is required?**	Pressure	**Which location on the cause effect chain is required?**	Impact midpoint
Select themes	Climate change, acidification, eutrophication	**Select themes**	Economy, energy and material use	**Select themes**	Toxicity
Results sub IK Impact assessment/ Theme selection	Endpoint Life Cycle Impact Assessment (LCIA) method		Input Output Analysis and Material Flow Analysis		Chemical Footprint method or Midpoint LCIA method,

Example 1		**Example 2**		**Example 3**	
What type of analysis is required?	Direction to target	**What type of analysis is required?**	Direction to target	**What type of analysis is required?**	Distance from target
				What type of sustainability goal is required?	A natural boundary
Which level of aggregation is required	Capitals	**Which level of aggregation is required**	Total	**Which level of aggregation is required**	Categories
What is the view on sustainability	Ecocentric	**What is the view on sustainability**	Weak		
Result sub IK aggregation	LCIA endpoint damage method	*Result sub IK aggregation*	A Multi Criteria Analysis (MCA) like weighted summation or Multi Attribute Value Theory (MAVT)	*Result sub IK aggregation*	Footprint method
Match of sub IKs→ method selection	Life Cycle Assessment with Endpoint LCIA method (e.g., ReCiPe)	**Match of sub IKs → method selection**	Material Flow Analysis and Input Output Analysis aggregated with MCA, e.g., MAVT	**Match of sub IKs → method selection**	Chemical pollution footprint method in combination with scenario building

3.3.2 Confronting sustainability assessments with the Identification Key

Next to illustration of hypothetical uses yielding vastly different methods for a single societal question (previous paragraph), we aimed to test the SA-IK for a suite of selected studies. We did that retrospectively. The goal of this analysis was to show whether and how the SA-IK may improve sustainability assessments on expected benefits mentioned above.

The case studies were selected by a two-step procedure: (1) a literature search from 2011 and beyond (Search engine: Scopus; search key: TITLE-ABS-KEY("Sustainability assessment" OR "Sustainability evaluation" OR "Sustainability performance") AND PUBYEAR > 2010); and (2) screening of titles and, in a second round, abstracts to sub-select the manuscripts that claim to describe a case study on sustainability assessment. Step 1 resulted in 1086 results and step 2 in a selection of 30 manuscripts with pertinent case studies. See the supplementary information at http://www.mdpi.com/2071-1050/7/3/2490#supplementary (09-03-2017) for an overview of the case studies. The introduction (goal/scope etc.), method description and results of the case studies were compared to the Domains and Criteria distinguished in the SA-IK. This analysis showed that the SA-IK potentially provides improvements in three directions: (a) more transparency in the link between the question and method selection, which is lacking or scanty available in most of the case studies; (b) more consistency between question and answer; and (c) more consistency in methodological design. These three conclusions are substantiated below.

3.3.2.1 Transparency on method selection

(a) Evaluation on Method selection. In 14 out of the 30 selected case studies, reasons for selecting a method are made explicit. Often one or two Criteria are mentioned as reason to select a method. As an example, considerations on "what is to be sustained" was mentioned as Criterion for method selection in 7 of the 30 case studies (Fig. 3.4). Fig. 3.4 shows that most case study descriptions revealed none or only little attention to method selection. This does not directly mean that methods were not carefully selected. It means that the relation between the question and the method selection was not explicitly described in the manuscripts (leaving room for variability in providing answers as illustrated in (Table 3.2). When criteria are given attention in the manuscripts they might also be taken into account in method selection, also when this relation is not explicitly described as such. For example, all case studies described the Object of study, but none of them explicitly brought this in relation with the method to be selected. This is visualized in Fig. 3.4. It shows the percentage of case studies in which the Criteria of the SA-IK are articulated (in dark grey) and the percentage of case studies in which the Criteria are explicitly brought into relation with method selection (in light grey).

The Criteria within the domain "System boundaries/Inventory" (object, spatial focus, temporal focus and life cycle or chain) are most frequently articulated in the case studies. For example, all case studies describe the Object under investigation. However, the temporal focus of the activity is discussed in less than 50% of the studies. In other words: it seems that in more than 50% of the case studies no explicit choices are made on the temporal focus of the activity under consideration. Of the criteria in the domain "Impact assessment and Theme selection", most (93%) case studies discuss what is to be sustained. However, the spatial and (especially) the temporal focus of the impacts assessed are often not discussed. Aggregation seems to play a role in 50% of the case studies, but the details of the aggregation, e.g., "Is a sustainability target required?" and "What is the view on sustainability?", are often not made explicit.

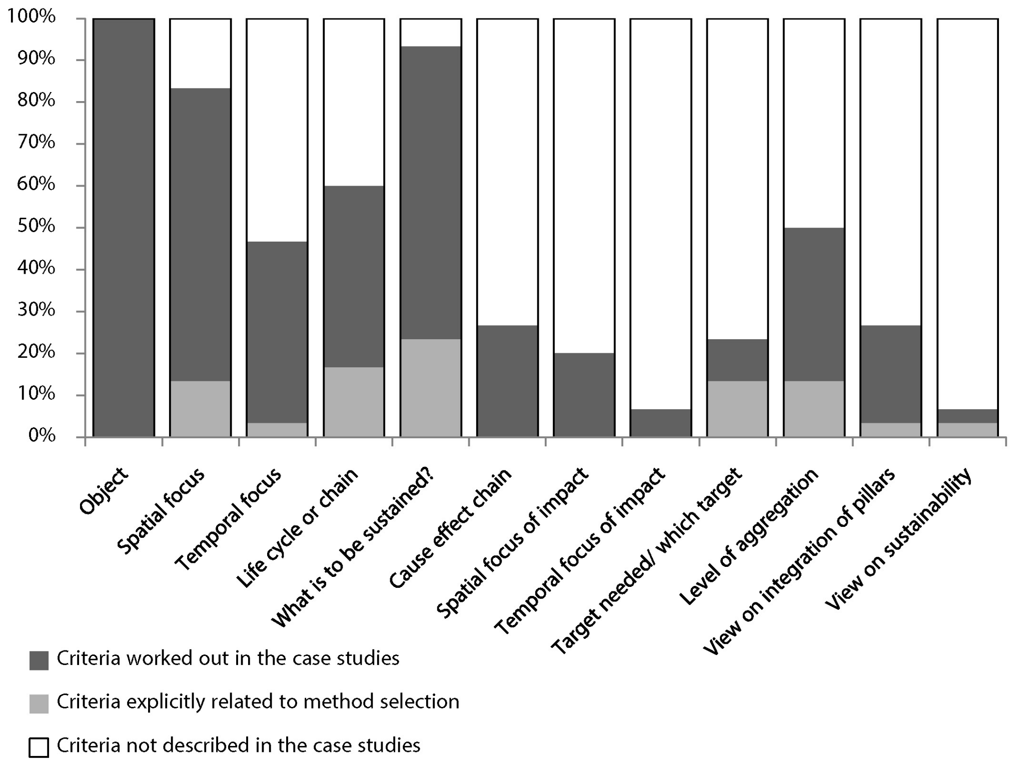


Fig. 3.4 Percentages of case studies in which Criteria were explicitly described (dark grey), percentage of case studies in which Criteria were explicitly coupled to method selection (light grey) and percentage of case studies in which consideration concerning this Criteria were not mentioned (blank).

3.3.2.2 Prevent inconsistencies between case descriptions and method selection

(b) Evaluation on consistency between question raised and answer provided. Lack of specificity in question articulation and in making tailored and transparent choices was hypothesized to lead to a mismatch between the question asked and the answers given. We analyzed in the same thirty studies whether the mismatch occurred. Four type of questions were distinguished: (1) determine how sustainable an object is; (2) compare alternative objects; (3) perform a hotspot analysis (which part of an object has the highest positive or negative impact on sustainability?); and (4) improve an object. A noticeable percentage of the case studies (6 out of 30, 20%) showed a mismatch between the type of question described in the case description and the answer provided as a result of the method chosen. This exemplifies that missing transparency in the step between "a question to be answered" and "the methodological design to answer the question" might lead to "a question not answered".

3.3.2.3 Prevent inconsistencies in methodological design

(c) Analysis on inconsistencies in methodological design. We hypothesized that lack of question articulation can lead to different specifications for different indicators within one assessment. For example, are themes quantified based on comparable system boundaries? More specifically: when a life cycle approach is thought to be important, is it incorporated for all indicators or only a selection? Life Cycle Analysis conceptually focusses on environmental impacts of a product, but social (Social-LCA) and economic (Life Cycle Costing) aspects can be performed from a life cycle perspective as well. However, some case studies that do include environmental, social and economic themes only perform a Life Cycle Analysis for the environmental ones and not for the social and economic themes. Another finding considers inconsistencies for the spatial scale. Often, within one study, world-wide environmental impacts are taken into account, whereas economic and social impacts are taken into account on the organizational or regional level (e.g., (Ibáñez-Forés et al., 2013; Mata et al., 2012; Traverso et al., 2012). Apparently, different capitals (People, Planet and Prosperity) tend to results in different spatial scopes. These inconsistencies were unexplained, such that SA-IK improvements can be gained here, though we also note that the observed inconsistencies are not necessarily wrong; we did not analyze the impacts of these inconsistencies on the results. However, in terms of question articulation the observed inconsistencies are remarkable. Logically, one view on the scope would be expected, e.g., one spatial focus, or otherwise an explanation for indicator-specific scope definition. Probably, some choices are not made explicit, leading to these inconsistencies in method design. Of the 30 case studies, 10 studies showed one or more of these types of inconsistency. With the SA-IK the preferred scope can be defined and compared with available indicators to choose from.

3.4 Discussion and conclusions

Method selection is a crucial step between a sustainability assessment question raised and an answer given. By choosing a method, one selects, or at least narrows down, the choice in system boundaries, themes and personal values. Therefore, method selection should be based on careful articulation of the original (societal) question. This is especially true in the field of sustainability assessments, with its manifold of available methods and interpretations of what sustainability or sustainable development actually means. Expanding on existing approaches, we observed unnecessary unclearness in the general literature on this subject, suggesting that a novel approach could be valuable for transparent, reproducible and valid method selection. The available organizing approaches (describing, characterizing and categorization) were analyzed to be largely supply-driven instead of demand-driven and appeared not to be capturing the dynamics and complexity that is needed for guiding the selection of methods for sustainability questions. This manuscript encompasses a plea for the design and use of a sustainability assessment identification key (SA-IK) as that next step. A SA-IK provides a modular approach that helps to structure question articulation and that leads to demand-based method selection. Functioning foreseen similar to a flora key for the taxonomy and identification of plant species, this identification key was designed to support using and categorizing large amounts of information, and to present the information needs in a step by step manner, which makes it manageable. The SA-IK helps those responsible for the sustainability question and the sustainability assessment, ideally supported by the relevant stakeholders, to select a method based on explicit (and as appropriate: improved) articulation of the question.

Deriving an identification key that makes sense has proven to be a worthwhile, but extensive process; a process that can take years. This will especially be true in a relative young field like sustainability assessment, where there is still much debate on terminology and data interpretation. On the other hand, as a side effect, developing the identification key might contribute to clarifications in this field by clarifying the relation between choices in definitions/interpretations and the consequences for the assessment.

Given existing sustainability assessment studies, as well as some principles for designing a taxonomical key, we provide a first attempt to design a Sustainability Assessment Identification Key (SA-IK), and use it in various ways. In other words, a first iteration (Fig. 3.2) of developing using and adjusting the SA-IK is provided. Based on examples, we have shown that although the design is incomplete and needs further development, the use of the SA-IK is supportive to:

I. guide and make explicit choices in method selection and design, revealing assumptions that remain hidden in many studies;
II. yield a better understanding of the question raised and how the question guides method selection
III. enable a more robust interpretation of the results, because the results can be placed in the context of methodological choices;
IV. producing eventually more transparent and reproducible assessments;

Furthermore, the SA-IK can provide insight in which type of question cannot yet be answered with the existing plethora of methods.

The proposed design is based on the observation that all sustainability assessments constitute three steps: (i) System boundaries/Inventory; (ii) Impact assessment/Theme selection; and (iii) Aggregation/Interpretation. Most Criteria for method selection found in literature have a role in only one of these three steps. The SA-IK itself consists of questions that can be answered for almost any problem definition, the answers guiding the assessors in various relevant directions. A single question can, depending on contextual aspects of the problem definition, result in different methodological choices, and a suite of questions can be analyzed systematically, such that the quality of sustainability assessments can be improved as compared to the current (reported) practices.

The SA-IK was not designed to provide answers to sustainability questions, but should serve in transparent and pertinent sustainability assessment method selection as such. Also, the SA-IK does not prescribe what sustainability assessment is and what sustainability assessment should encompass, but is designed to make these choices case specific, with all the relevant stakeholders, in a transparent reproducible and explicit way. The key reveals the consequences of choices for method selection, but does not prescribe these choices. Efforts to find consensus regarding the definition of sustainability and sustainability assessment exist, for example the development of the Bellagio STAMP principles (Pintér et al., 2012), as do attempts to describe the ideal sustainability assessment method (Sala et al., 2012b). These were taken into consideration for the first design of the SA-IK.

Thirty case studies on sustainability assessment that were recently published in literature were evaluated based on the SA-IK. The analyses showed that using the SA-IK makes many hidden choices explicit, but also reveals inconsistencies, which would have been avoided had the SA-IK been used. In 6 of the 30 case studies, limited question articulation appears to have led to a mismatch between the type of questions asked and the type of answer provided and in 10 of the 30 case studies to an inconsistent method design. Hence, the SA-IK use in its current format is potentially helpful in improving sustainability assessments.

We expect that the iterative process of using, discussing and further developing the identification key will at least lead to more transparency in method selection and potentially also to a better match between questions asked and answers provided.

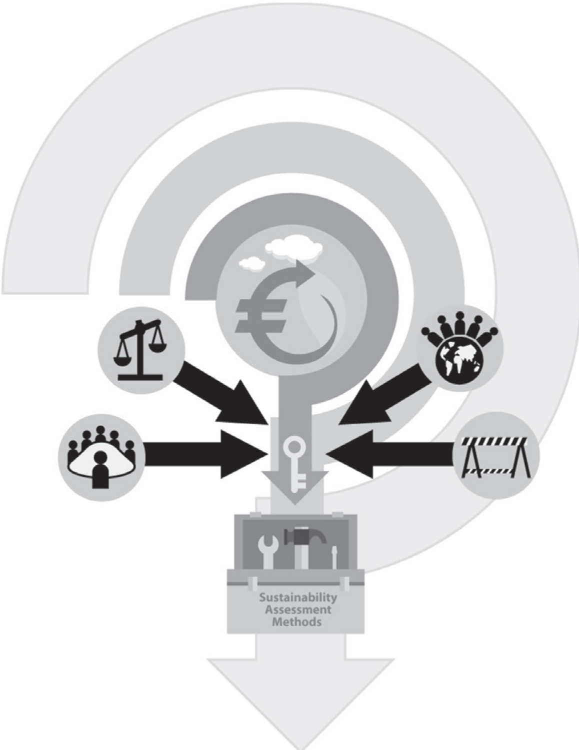
3.5 Acknowledgments

Funding by the Strategic Research Program of RIVM is gratefully acknowledged (project number S/607020).

4 Method selection for sustainability assessments: the case of recovery of resources from waste water

Authors: Michiel C. Zijp, Susanne L. Waaijers-Van der Loop, Reinout Heijungs, Martijn L.M. Broeren, Ruud Peeters, Arjen F. Van Nieuwenhuijzen, Li Shen, Evelyn H.W. Heugens, Leo Posthuma

Published in: Journal of Environmental Management (2017), 197, 221-230



Abstract

Sustainability assessments provide scientific support in decision procedures towards sustainable solutions. However, in order to contribute in identifying and choosing sustainable solutions, the sustainability assessment has to fit the decision context. Two complicating factors exist. First, different stakeholders tend to have different views on what a sustainability assessment should encompass. Second, a plethora of sustainability assessment methods exist, due to the multi-dimensional characteristic of the concept. Different methods provide other representations of sustainability. Based on a literature review, we present a protocol to facilitate method selection together with stakeholders. The protocol guides the exploration of i) the decision context, ii) the different views of stakeholders and iii) the selection of pertinent assessment methods. In addition, we present an online tool for method selection. This tool identifies assessment methods that meet the specifications obtained with the protocol, and currently contains characteristics of 30 sustainability assessment methods. The utility of the protocol and the tool are tested in a case study on the recovery of resources from domestic waste water. In several iterations, a combination of methods was selected, followed by execution of the selected sustainability assessment methods. The assessment results can be used in the first phase of the decision procedure that leads to a strategic choice for sustainable resource recovery from waste water in the Netherlands.

4.1 Introduction

Transition towards a circular economy has been proposed as one of the solutions for a future that supports the growing world population and welfare per capita within the environmental and social boundaries of our planet (Ellen MacArthur Foundation, 2013). Initiatives towards realizing a circular economy can be found from global to local levels (Bocken et al., 2016; European Commission, 2016a; Geng et al., 2013; Linder and Williander, 2015; Ministerie van Infrastructuur en Milieu, 2016; Municipality Utrecht, 2015; UN, 2015a). The transition process needs be supported by insights from different disciplines with respect to the economic, environmental and social costs and benefits, amongst which trade-offs may occur. In addition, decision makers have to deal with uncertainties and unknowns that are characteristic of investing in new business models (Linder and Williander, 2015), and with different stakeholders views on the current situation, the desired solution and on what sustainable choices should encompass (Zijp et al., 2015). The selection of sustainable solutions for a resource-efficient economy is a wicked problem sensu Rittel and Webber (1973).

An example of the need for such strategic choices in the realms of circularity is the recovery of resources from domestic waste water. We selected this as a case study to design and test an approach to support decision-making with a sustainability assessment. We applied a solution focused sustainability assessment framework (Zijp et al., 2016), with specific focus on the translation of the sustainability question and its context into sustainability analysis methods selection (Zijp et al., 2015).

Currently, the use of waste water flows as a potentially valuable resource are evaluated in various pilot projects in the Netherlands and elsewhere. In order to invest in full-scale operations, water system managers need to make strategic choices. The choices involve technical issues (e.g. different solutions for resource recovery from waste streams are possible but can be mutually exclusive), political issues (e.g. the focus on climate change draws organizations towards low-energy cost solutions without considering the biomass value pyramid (Gavrilescu, 2014)), many unknowns (e.g. what will be the future quality of waste water) and many stakeholders.

A sustainability assessment (SA) provides scientific support in the decision-making for selecting amongst competing sustainability-enhancing technologies. Its outcomes can be utilized in a decision-making process that is solution focused, participative, iterative and transparent in its definition of sustainability (Zijp et al., 2016). However, many SA methods can be utilized and the question arises: which (set of) SA methods is most suitable for the evaluation of a specific situation? In practice, assessment methods are often selected by an expert, with poor question articulation and with limited inclusion of stakeholders' views on sustainability (Zijp et al., 2015). This approach can lead results that are incomplete in their coverage of the sustainability metrics of relevance, and may furthermore not be supported by the stakeholders and decision makers, so that they are consequently of limited practical influence in the decision context.

In order to support the consistency and utility of SA, Zijp et al. (2015) proposed the idea of a sustainability assessment identification key, to identify case-specific requirements for a SA and use these requirements to make selections amongst the available SA methods. The key supports a transparent and well-considered choice for an SA method or combination of methods. Furthermore, it specifies what can and cannot be expected from the assessment. Since its publication, the proposed SA-methods identification key has been applied to studies that report transparently on method selection (e.g., Moreira et al. (2015)), but not yet in its inverse application: to first determine the specifications of a transition plan, and then select a method. This process, of setting the requirements for an SA, is further referred to as 'question articulation'. Ideally, question articulation is performed together with the stakeholders (Harder, 2015). Firstly, because every stakeholder can contribute to the question articulation with knowledge on potential problems and solutions from a diversity of points of view (Zijp et al., 2016); and secondly, because transparent and participative processes enhance the trust in the outcomes and in the choices based on the outcomes (Lind et al., 1990; Lind and Tyler, 1988; Tyler and Lind, 1992).

In practice, operating the SA identification key for use in a process of finding solutions for wicked problems requires insight in what is required (the demand), knowledge on which methods are available to choose from (the supply) and a procedure to select the most suitable (combination of) available methods based on the question articulation and the inventory of available methods. This paper reviews a non-limitative set of existing SA-methods, and thereupon provides a transparent way to select a method for sustainability assessment from those, based on participative question articulation, in the context of its case-specific decision process. In detail, the aims are to:

1. provide a protocol for sustainability question articulation (the demand-side of method selection);
2. review available sustainability assessment methods for products (the supply-side of method selection)
3. provide a tool that matches the demand- and supply side of method selection; and
4. evaluate and illustrate the applicability of the protocol and the tool to the early strategic decision stage of a wicked problem case study: the recovery of resources from waste water in the Netherlands.

This paper is structured as follows: in section 4.2 the approach to draft the protocol (§4.2.1), perform the review and design the tool (§4.2.2) are explained, and the case study is introduced (§4.2.3). Then, in section 4.3, the results are described: first the protocol (§4.3.1), then the results of the review of SA methods and the tool in which the review results are translated (§4.3.2) and finally the results of applying the protocol and the tool on the case study (§4.3.3), including the actual sustainability assessment. The discussion on the approach, the case study results and the meaning of the results for application of the protocol and tool in other studies are presented in section 4.4. Finally, main conclusions are summarized in section 4.5.

4.2 Research approach

A protocol (§4.2.1) and a tool (§4.2.2) for question articulation and method selection were developed, based on a review of currently available SA methods and participation approaches. The applicability of the protocol and tool was tested in a case study (§4.2.3). The Annex in §4.7.1 defines the terms used in this paper, acknowledging that different literature sources use different terminologies for similar matters.

4.2.1 Protocol for question articulation

The goal of the protocol is to help specifying the sustainability question(s) and the(ir) context such that the transition problem can be translated into sustainability metrics that should be part of the SA. The protocol supports the systematic exploration of the problem definition and its solution scenarios, and merges the possible diversity of inputs from different stakeholders into requirements for the method selection. For example, the protocol helps to sub-select pertinent themes (such as climate change, economic performance and social equality, Fig. 4.2), as sustainability assessments do not have to include all themes of sustainability (Laniak et al., 2013). In complex situations this may be an iterative process, involving different groups of stakeholders, e.g. first with the internal stakeholders that are involved in the design of solution scenarios and then with external stakeholders. The protocol was aligned with approaches described in the literature on stakeholder participation, such as https://www.irgc.org/stakeholder (visited 02-11-2016), and expertise of the authors on transparent decision-making.

4.2.2 Tool with available sustainability assessment methods

The goal of the SA-method selection tool is to transparently translate the question articulation provided by the protocol into a selection of methods pertinent to the question.

The tool was designed by making a database of existing SA methods. An inventory of available sustainability assessment methods was drafted based on expertise within the consortium, an iterative search in google scholar using the keywords: "Sustainability assessment" AND "Products" AND/OR "Tool", AND/OR "Method", AND/OR "Methodology", AND/OR "Approach" and input from a group of experts in the field of sustainability assessments. Methods were selected that:

- assess either products or feedstock (thus excluding methods that assess the sustainability of organizations, such as the Dow Jones Sustainability Index); or
- are applied at least once; and
- are transparently described in accessible sources.

This approach was meant to result in a non-limitative overview of methods that are available for practical application. Hence, this overview (and the tool) can be expanded.

The selected methods were analyzed and categorized using the sustainability assessment method selection criteria found in literature and reported in Zijp et al. (2015). It was scored, for example, which themes are covered by one or more of the indicators of a method, and what spatial and temporal scales covered by the method are? The list of criteria and their definitions are reported in the Annexes of this chapter (§4.7.2., Table S4.1).

The tool was designed such that it links the specifications of the question articulation (protocol) with the criteria of the reviewed methods. The protocol and tool can be combined or used stand-alone, depending on the expertise and problem at hand.

There may not always be a particular method available to fit the question. Therefore, like the protocol, the tool supports an iterative working process, changing the method requirements until a 'best method available selection' can be made. Although presentation of the SA-results, e.g. by aggregation of various indicators in an overall score of each solutions, is an important aspect of method selection (Brewer and Stern, 2005; Laniak et al., 2013; Özdemir et al., 2011; Zijp et al., 2015), it was not made part of the protocol and tool.

4.2.3 The case study

A case study was performed to evaluate and illustrate the utility of the protocol and tool for assessing sustainability in an early stage decision context. The case concerns the opportunities for recovery of resources from waste water, evaluated by the Energy and Resource Factory (ERF). ERF is a consortium of Dutch water boards' innovation and sustainable development strategists. It explores solutions to treat domestic waste water as a resource instead of a waste-stream.

Presently, various solutions to extract resources from waste water are operational at different water boards in the Netherlands (STOWA, 2016). Only the recovery of energy has reached a matured implementation stage and has full-scale applications. Upscaling of pilot installations for the recovery of other resources to full-scale operations requires significant investments. Also, different solutions for resource recovery can be mutually exclusive. Alternative options imply different types and magnitudes of impacts across different SA themes. Therefore, strategic choices have to be made on the combination of recovered resources that yield the best sustainability performance. This case study is an early-stage exploration of the potential sustainability aspects of different strategies to recover resources from domestic waste water. That is, the assessment should support the first iteration of the decision procedure towards strategic choices for large investments: how to interpret and obtain sustainable development and what are the differences between the resource recovery and utilization pathways from a sustainable development point of view?

The following five waste water resource recovery and utilization options are considered in this study and are presently under investigation and/or in operation by ERF water boards:

- Alginate is a substance that is available in algae and seaweeds, but can also be produced by bacteria. Alginate is used by various industries for its adhesive qualities.
- Biogas is currently produced from sludge in waste water treatment processes as an energy resource. It is included in this study as comparison to the new resource recovery solutions.
- Cellulose are fibers, mainly from toilet paper, that can be recovered from the waste water and

reused in several applications, of which we include fermentation (biogas), pellets for incineration to produce bioenergy, replacement of cellulose in construction (asphalt, concrete, isolation) and application in the production of carton or paper.

- Polyhydroxyalkanoate (PHA) is a polymer that can serve as basis for bioplastic products. It can be produced by feeding sludge biomass fatty acids that can be extracted from waste water.
- Phosphorus is a nutrient that can be recovered from waste water as struvite (at the waste water treatment plant, WWTP, which is further referred to as decentral) or as phosphor (after incineration of the sludge, which we will refer to as central) and used as fertilizer.

A general flowchart of a waste water treatment plant (WWTP) detailing process steps at which the above resources can potentially be recovered is provided in Fig. S4.2 (annex §.4.7.6).

4.3 Results

4.3.1 Protocol for question articulation

The protocol developed for question articulation (table S4.3 in the annexes §4.7.3) consists of three parts.

The first part explores the context in which the participants of the SA-process operate and whether (and how) sustainable development can be related to this context. For example: what are the major challenges the company faces? Different internal and external stakeholders are requested to provide their views on what the challenges are and these views are then clustered by the participants (and with that discussed). The resulting overview of challenges is then discussed in the light of sustainable development goals. This first step is important because i) it provides the context of the SA from different perspectives; ii) it reveals the relation of the SA with the actual working process of the whole system of participants; iii) it provides insight for all participants in each of the stakeholders' view on sustainable development, which, in return, leads to improved understanding of each other's input in the process; and iv) it creates insight in what is expected of the SA, in the format of eventual decision-support information. The process was developed such that interaction between the participants is stimulated, while securing individual input from every participant.

The second part of the protocol is the question articulation and consists of a series of questions to specify the sustainability question such that method selection can be transparent and based on the stakeholder preferences (table S4.4, §4.7.3). During this phase participants discuss what they think is important and motivate why it is important.

The third part of the protocol is the discussion of the question-articulation based steps towards method selection and the design of the further process. The resulting action plan documents the actions required to start the sustainability assessment. This can be for example a flow chart of the data that needs to be gathered, an iteration of the question articulation with other stakeholders, or a more detailed exploration of a selection of methods.

4.3.2 Tool for method selection

The review search for available SA methods resulted in 30 methods (Fig. 4.1) that fitted the search profile (§4.3.2.2). Although non-limitative, the list provides an overview on the diversity in methods that are available. The results of reshaping the method analysis results into the design of an operational SA-methods selection tool are summarized below. It should be noted that the results below are observations and not assessments. For example, the fact that method A covers more themes than method B does not make method A better than B. The final goal of the tool is to match SA-questions to available (combinations of) method(s), and not to rank or validate methods. Evidently, identified methods should be scrutinized for quality aspects and validity.

4.3.2.1 ystem boundaries covered by the methods

The methods differ in their system boundaries (Fig. 4.1). Of the 30 methods, 18 methods apply life cycle thinking on all themes covered by the method, either cradle to gate or cradle to grave. Five methods hold a combination of themes that are, or are not assessed with life cycle thinking. The other seven methods do not take life cycle thinking into account. The spatial focus of the methods ranges from anywhere (generic metrics and data are used) to local (assessment of a site-specific situation). Again, within methods the spatial focus can differ between themes. Furthermore, methods that include life cycle thinking often use a combination of site-specific data (foreground data) and regional-average and/or generic data (background data). This is common practice in life cycle assessments (EC-JRC, 2010). With regard to the temporal focus, we observed that although most methods could be suitable for a retrospective or prospective analysis, the default settings and primary use are snapshots. That is: the assessment is based on data and knowledge from a recent time-lock (Harder, 2015).Two methods are explicitly designed to look forward (prospective). Finally, methods are available for all technical development- and implementation stages, with relatively the lowest coverage for the development stage.

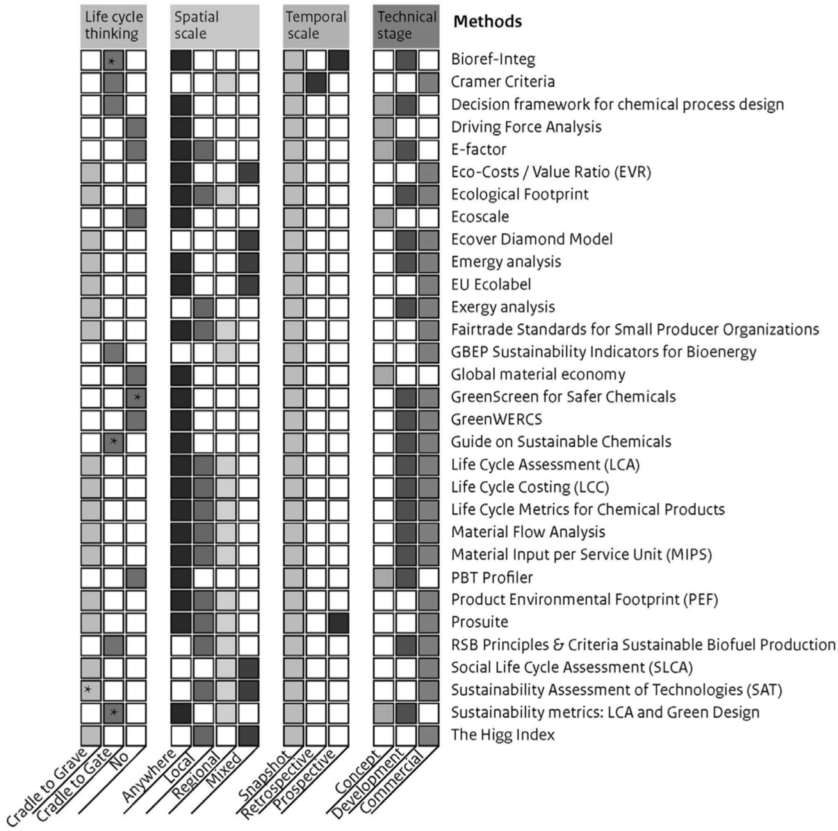


Fig. 4.1 Overview of the system boundaries of the methods collated in the review of this study; for references of the methods see the annexes §4.7.4, table S4.5. The asterisk (*) shows that for this method life cycle perspective is taken into account for part of the themes (Bioref-Integ and Guide on Sustainable Chemicals only qualitative; Greenscreen only for the theme "Biodegradation"; SAT only for the themes "Climate Change" and "Economic Performance"; Sustainability metrics only for the theme "Energy efficiency").

4.3.2.2 Themes covered by the methods

Methods differ in their coverage of sustainability themes. Furthermore, different methods use different names for comparable themes. In order to provide an overview of theme coverage and to make this applicable for the method selection tool, we harmonized and structured the themes. The themes covered by the methods were arranged under six 'areas of protection', so called capitals: economic welfare, environmental and biological quality, human health, social welfare, resources and technological welfare. The choice for these capitals was based on existing frameworks (Blok et al., 2013; UN, 2015a) and expert judgment (see for definitions Table S4.2 in the annexes to this chapter, §4.7.2). Several qualifications (e.g. transparency of the assessed techniques) were encountered in SA methods that could not be linked to the capitals, yet such meta-information can be valuable to judge. These themes were grouped under the capital Technological Welfare. Based on the literature, we distinguish the following themes within Technological welfare: innovation (the potential of an innovation to open up new markets and product applications can in some situations be an argument to accept uncertainties in its long term environmental benefits, economic gain and human health improvements); transparency (is the solution and its assessment transparent, are participants willing to share information about the technology under investigation); feasibility (what happens when a solution is introduced at large scale: are enough feedstocks available); and flexibility (can the solution be adapted, e.g. replacement of resources). The themes covered per Capital (area of protection) are visualized in Fig. 4.2.

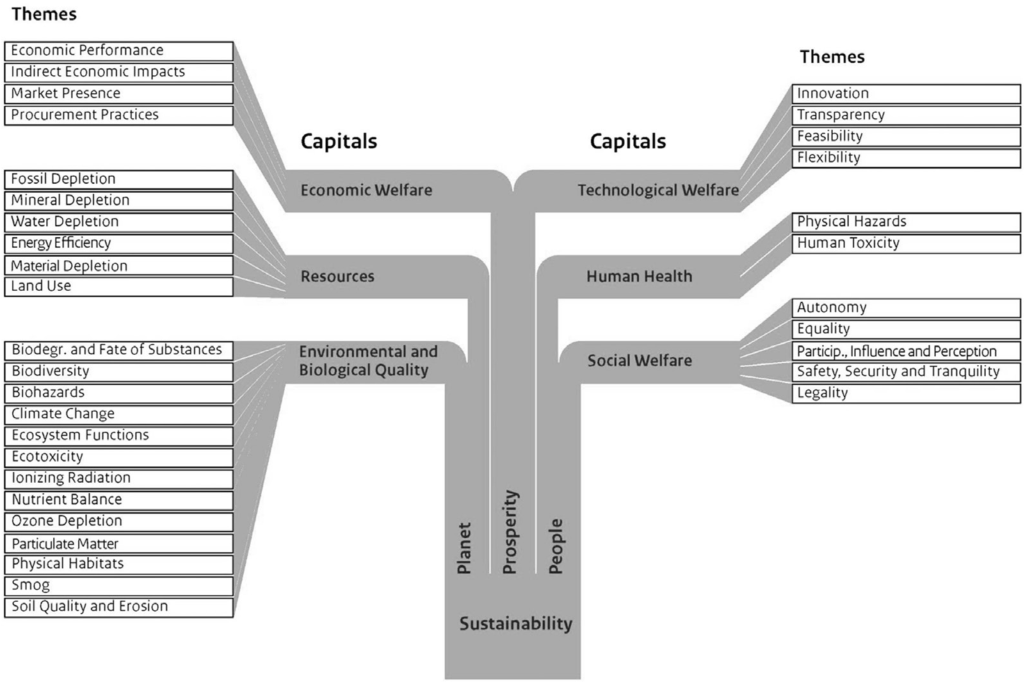


Fig. 4.2 Overview of the capitals (areas of protection) and themes distinguished by the different methods.

4

The capitals Environmental and Biological quality, Resources and Human Health are represented in most of the methods (~75%), whereas the other capitals are less covered, with the lowest coverage for the capital Technical Welfare (3 of the 30 methods, Fig. S4.1 in the annexes to this chapter, §4.7.5). Two of the methods cover all capitals, while five methods focus on only one capital. An overview of the themes covered per method can be found at www.sustainabilitymethod.com (last visited at 21-02-2017).

4.3.2.3 Online version of tool

A tool for supporting utilization of the collated knowledge on method characteristics was designed in the format of a decision tree, based on the protocol steps, and implemented on a website (www.sustainabilitymethod.com). The tool works with filters for the system boundaries and scores for the theme coverage. That is, the required system boundaries and themes can be chosen in the tool as selection-relevant criteria. The use of the tool results in an overview of the possible (set of) method(s) that fit the selected requisites of a case. The selected method(s) need to be checked prior to utilization in the SA, to ascertain that methodological details are fit to address the problem. If not, the identified method is a 'most-similar' methodological approach, on the basis of which a tailored SA can be designed.

4.3.3 Process aspects of the case study

The case study has been executed with focus on the process steps (the context exploration and the method selection as part of the solution focused SA framework) and on a judgement of the outcomes for the resource recovery scenarios. The protocol was applied in two workshops with members of the water boards. Part one and two (the context and question articulation) were covered in the first workshop, and part three (the method selection and action plan) in the second workshop.

4.3.3.1 Context exploration

Evaluation of the context revealed that ERF members are intrinsically motivated to work on recovery of resources from wastewater, because the majority of the members expressed the believe that a circular economy is more sustainable than a linear economy. As an organisation, ERF expressed the hope that an organized sustainability assessment process and analysis can support the process towards a shared vision on the role of the water boards for a circular economy, and in preparing the strategic choices for one or a combination of resources and techniques to focus on. ERF stated that they want sustainable development as leading principle for technical innovations in the future, both in an environmental and in a socio-economic way. That is, environmental in the sense that environmental impacts are reduced compared to the present situation and socio-economic in the sense that the resulting products/techniques have a healthy economic performance in which, further, innovations are stimulated.

4.3.3.2 Question articulation

The context exploration (table S4.3, annex §4.7.3) was followed by the question articulation (table S4.4 in the same annex) and resulted in the following specification of the sustainability question, and thus in requirements for the sustainability assessment (method selection):

- The question focuses on products as object. The focus of the ERF is on products that are based on recovered resources from wastewater and are more sustainable than their alternatives. The goal is not recovery of resources from waste water per se, but a net more sustainable practice.
- The chosen life cycle perspective is from the waste water flow (cradle) to the use (including

disposal) of the products that are based on the recovered resource (grave).

- The spatial focus of the SA is national. Thus, the potential of all WWTPs in the Netherlands should be taken into account. The assessment has thus to result in conclusions on the macro level: taking into account actual flows of resources. The life cycle of the products that are based on the recovered resources can contain activities that are outside the Netherlands, such as the production of chemicals that are required for an extraction procedure. These activities are taken into account.
- The temporal focus of the SA is the present. Strategic choices must be taken the coming five years.
- The technical stage of the techniques under consideration range from development to commercial stage, i.e., technology readiness level 3 - 9 (EC, 2014b).
- Finally, the themes that the participants mentioned to be important were: climate change; economic performance; energy efficiency; feasibility; fossil depletion; human toxicity; innovation; land use; legality; material efficiency; mineral depletion; nutrient balance (eutrophication); perception; physical habitat; physical hazards; safety, security and tranquility; salinization; water depletion. The theme perception covers the interests of and barriers for different stakeholders to support and use resources recovered from waste waters (Liang and Van Dijk, 2016). The themes physical hazards, safety and perception are to be understood in relation to the perception on potential microbiological hazards of resources recovered from waste water.

4.3.3.3 Method selection

Of the 30 existing SA-methods, 10 appeared to fit the system boundaries chosen at the workshops. Of those 10 methods, Prosuite (Blok et al., 2013) covers most of the themes mentioned (12 of the 16 themes, Table 4.1). One theme mentioned during the workshop, salinization, is not yet part of the overview of themes that was created based on the review (Fig. 4.2). Although salinization relate to the themes climate change and water depletion (main causes of salinization) the impact of salinization due to emissions, e.g. from WWTPs, is not covered in any of the reviewed methods. Other themes fully overlap. For example: material efficiency is not an explicit theme (or in LCA terms: impact category) in LCA, but is implicitly taken into account in the assessment: higher material efficiency results in less mineral depletion, fossil resources depletion and emissions (e.g. climate change). This type of indirect coverage of themes by methods is indicated with an 'i' in Table 4.1.

Prosuite covered most of the selected themes (12 of the 16). However, data for micro-economic analysis as described in Prosuite was not available for most of the solutions for reasons of absence of data and data confidentiality. Furthermore, the weighing applied in Prosuite to aggregate results was not required for this case study.

The resulting set of potentially selected SA methods was evaluated in interaction with the ERF. As a result a combination of methods was chosen for the final SA that was selected: a life cycle assessment approach (LCA, for the themes under Resources, Environmental quality and Human health), a literature review and expert elicitation (for Economic welfare and Social welfare) and an evaluation of the technical readiness level (EC, 2014b) (TRL, for Technical welfare). This combination of methods was applied to summarize the sustainability differences between the current practice of WWTPs, including retrieval of biogas, and the alternatives described in section 4.2.3.

Table 4.1 The ten methods that fit the system boundary specification of the case study, and their theme coverage (chosen by the ERF in the workshop). An x indicates direct coverage; an i indicates that the theme is covered indirectly via an indicator used for another theme in the method.

Method		a	b	c	d	e	f	g	h	j	k
Economic welfare	Economic performance	x			x	x					x
Resources	Energy efficiency	i	i	i	i						
	Fossil depletion	x	x	x	x		x		x		
	Land use	x	x	x	x					x	
	Material efficiency	i	i	i	i		x		x		
	Mineral depletion	x	x	x	x		x				
	Water depletion	x	x	x	x		x				
Environmental and biological quality	Climate change	x	x	x	x	x				x	
	Nutrient balance	x	x	x	x						
	Physical habitat	i	i	i						i	
	Salinization										
Technical welfare	Feasability										
	Innovation										
Human Health	Human toxicity	x	x	x	x						
Social welfare	Legality					x		x			
	Perception of biological safety	x				x		x			

a = Prosuite; b=Life cycle metrics for chemical products; c=Product environmental footprint; d=Eco-cost value ratio; e=Sustainability assessment of technologies; f=Material input per unit service; g=Social LCA; h=Emergy analysis; j=Ecological footprint; k=Life cycle costing

4.3.4 Sustainability assessment outcomes of the case study

The first goal of the assessment was to gain a general overview on how the five potential resource recovery solutions differ in their impact on the selected themes compared to the present situation: are the solutions for a more circular business model more sustainable the present business model? The second goal was to compare the solutions among each other. Which solutions should be considered by the ERF for large scale investments? The results are summarized in Table 4.2 and discussed below.

4.3.4.1 Economic welfare

The economic performance was assessed using existing market exploration studies complemented with expert judgment. The exploration of existing studies (references in the annexes §4.7.7) resulted in insights in expected production costs and market values for the different solutions.

As shown in Table 4.2, all solutions are potentially economically feasible (scores ≥ 1). Economic performance appeared to depend on the WWTP-specific features. Furthermore, it depends on the temporal scope of the assessment. For example, struvite production is not competitive (the production costs are higher than the market value) on the short term, but when taking into account that its recovery would imply reduction of maintenance of the WWTP in the long term it was expected to be economically neutral. The business case with respect to the economical welfare seemed promising for alginate and most of the cellulose applications. The application of cellulose as source for carton or paper seemed, however, less beneficial due to the costs of an extra hygiene step, which was discussed as a need for this application. For PHA, just like struvite, the business case depends on location-specific aspects, but it appeared promising enough for further investigation.

4.3.4.2 Resource depletion

The themes fossil depletion, mineral depletion, land use and water depletion were quantified using life cycle assessment (LCA) (Van Nieuwenhuijzen et al., 2016). The life cycle that was studied included the retrieval of resources from waste water, the production of products from those resources and the use and disposal of these products. In order to be able to compare the different solutions, the functional unit for the LCA was set at 100.000 population equivalents of inflow of a medium-consistency type of domestic waste water. The functional unit is consistent with the function of WWTPs: treating domestic waste water. The assessment is aimed to quantify the difference in impacts when implementing the solutions compared to normal operations, i.e. it covers the additional interventions and operational changes. Furthermore, the impacts are calculated compared to the present situation at the WWTPs, including biogas production. Thus, when a solution for recovery of a resource results in less biogas retrieval compared to the present situation this was accounted for in the results. Finally, when a product replaces another product this is accounted for in the LCA, e.g. using struvite instead of another fertilizer. Further details of the LCA were reported in Van Nieuwenhuijzen et al. (2016) and STOWA (2016).

The results show that the novel resource management pathways often resulted in a reduction in the depletion of resources (fossils, minerals, land, water). In more detail, decentral phosphorous recovery results in more reduction of resource depletion than central recovery. PHA production resulted in a slight increase in mineral and fossil depletion, but in less land use and water depletion effects. Cellulose recovery and application in construction and paper industry resulted in extra fossil depletion compared to the current situation, due to an extra preparation step prior to this use in the paper and construction industry. On the other hand, these solutions result in high reductions of mineral depletion, land use and water depletion. Use of cellulose for biogas and bioenergy are solutions at the bottom of the biomass value pyramid (Gavrilescu, 2014). However, they do result in reduced use of fossil fuels. Finally, alginate production shows beneficial results for all resource themes.

4.3.4.3 Environmental and biological quality

The themes climate change and nutrient balance were assessed using LCA, as described above for Resource depletion. Again, decentral struvite production scores better than central production. The only solution that has higher impacts on environmental quality than the reference situation is the use of recovered cellulose for paper production. This is a result of the energy use for the extra hygiene step that is thought to be required for use in the paper industry. This hygiene step is not required for use of gained cellulose in constructions, which is why that solution scores positive (less impact) compared to the reference situation.

4.3.4.4 Technical welfare

Feasibility and innovation were assessed using the TRL method, which provides a score to technologies based on expert elicitation. The scores range from 1 (basic principles observed) to 9 (actual system proven in operational environment) (EC, 2014b). Details of the assessment can be found in STOWA (2016). At the time of the SA, the solutions were characterized by differences in the technical stage. Recovery of biogas is more or less the standard in present-day waste water treatment, while products from alginate or cellulose are still in a pre-mature stage. Of all solutions, struvite and the use of cellulose for biogas showed the highest technological readiness level, i.e. successful mission operation, and alginate production from waste water the lowest: validation in lab. PHA and cellulose for paper and for construction scored a TRL of 5 (technology validated in relevant environment) and cellulose for bio-energy a TRL of 7 (system prototype demonstration in operational environment). This score is reported separately from the other themes (not in Table 4.2), because a low technical readiness level is not necessarily a negative feature.

4.3.4.5 Human health

Human toxicity was assessed using LCA, as described above under Resource depletion. Only PHA production results in emissions of toxic substances to the environment that may cause higher impacts on human health than the emissions in the reference situation.

4.3.4.6 Social welfare

The impacts of regulation (legality) and risk perception of stakeholders on the safety of resources from waste water on the market chances of the solutions were assessed using expert elicitation, following the guidelines provided by Gaasbeek and Meijer (2013) (for detailed information see the Annexes 4.7.8). Next to market influence, the relevance of risk perception was assessed compared to perception of the market on the supply security, on the costs and on the sustainability of the solution. According to the experts, risk perception is less important than the other issues for struvite and application of cellulose for biogas, bioenergy and in construction. For the other solutions, risk perception was equally important (PHA and Alginate) or more important (application of cellulose in paper and carton industry) than the perception of the market on the supply security, on the costs and on the sustainability of the solution. In general, risk perception is seen by the experts as a problem that can be solved on the short term. One product can pave the way for other products, as long as there are no large incidents or calamities. It was further noted that not everybody has to feel comfortable with the products to be successful. A relatively small group of customers might be enough for a successful circular solution. There is an exception for the paper industry using waste water resources. This relates to a negative risk perception (hygienic aspects) and because experts estimate the volume of this route to be too low for practical large-scale application. Legality is interpreted as the legally attributed status of the resources, that is: whether it is formally assigned to be either a waste or a product. This assignment is considered to present a serious challenge for the success of resource use, which needs to be solved in order to become successful. Legality scores more or less alike for all solutions, except for bioenergy from cellulose. The other solutions require a change in status from waste to resource, which was remarked to be preferably accepted at an international level (e.g., the EU) in order to support a successful implementation. More details on the expert elicitation and its results can be found in the Annexes to chapter (§4.7.8).

4.3.4.7 Overall results

All analyses steps resulted in the classified outcome scales, but were also characterized by uncertainty (not shown). This is partly a result of the intrinsic uncertainty in the applied methods, but merely due to the premature development and application phase for most of the solutions which inhibits an accurate estimation (Van Nieuwenhuijzen et al., 2016). This is typical for a SA in an early stage of scenario definition and evaluation. Nonetheless, the SA outcomes were considered by the ERF to be useful in the decision-making procedure as follows. First, the recovery of resources from waste water indeed appeared to lead to a reductions in environmental impacts and resource depletion. Secondly, it was shown that there is not one solution that scores best on all selected themes. Hence, choices will require evaluation of trade-offs and weighing of the results. Thirdly, based on the results, one of the solutions, paper or carton from cellulose, could be decided to be left out of further assessments and technological development, because implementation would be prohibited in relation to negative safety perceptions of the paper industry towards the product (safety and volume). Finally, continuing the decision procedure for an investment in one or a combination of resources requires a new iteration with the protocol, including stakeholders in the method selection process. The SA can then be broadened by analyzing WWTP-specific possible combinations of resource recovery (De Fooij, 2015; Van Nieuwenhuijzen et al., 2016) and take into account actual yield estimates at the different WWTP in the Netherlands (Van Nieuwenhuijzen et al., 2016).

Table 4.2 Results of the sustainability assessment. The economic welfare and social welfare themes are expressed in categories. The higher the score the better the performance of the solution. Market values lower than the range of production costs indicated a negative economic performance (implementation not likely, category 0), market values above the range of production costs indicated a positive performance (most likely, category 3) and overlapping ranges of costs and market values were an indication of possible (category 1) or likely positive economic performance (category 2). The resources, environmental and biological quality and human health themes are expressed in their LCA units and in comparison with the current situation (waste water treatment with biogas production). Hence, a negative value indicates a benefit compared to the reference situation and positive values and vice versa.

Capital	Theme	Struvite Central	Struvite Decentral	PHA	Cellulose Biogas	Cellulose Bioenergy	Cellulose Paper	Cellulose Construction	Alginate
Economic welfare	Economic performance	1	1	1.5	3	3	1	3	3
Resources	Fossil depletion (kg Oil eq.)	$-1.1 \cdot 10^4$	$-8.2 \cdot 10^4$	$1.0 \cdot 10^4$	$-1.7 \cdot 10^5$	$-1.4 \cdot 10^5$	$6.3 \cdot 10^4$	$2.2 \cdot 10^4$	$-1.5 \cdot 10^5$
	Land use (m^2)	$-9.0 \cdot 10^2$	$-1.1 \cdot 10^4$	$-2.4 \cdot 10^5$	$-2.5 \cdot 10^4$	$-1.9 \cdot 10^4$	$-1.3 \cdot 10^6$	$-1.3 \cdot 10^6$	$-1.1 \cdot 10^4$
	Mineral depletion (kg Fe eq.)	$-8.5 \cdot 10^2$	$-3.2 \cdot 10^4$	$2.4 \cdot 10^3$	$-6.1 \cdot 10^3$	$-5.6 \cdot 10^3$	$-1.1 \cdot 10^4$	$-1.2 \cdot 10^4$	$-6.0 \cdot 10^4$
	Water depletion (m^3)	$2.3 \cdot 10^2$	$-8.5 \cdot 10^3$	$-1.4 \cdot 10^4$	$-1.8 \cdot 10^3$	$-1.6 \cdot 10^3$	$-6.2 \cdot 10^3$	$-6.4 \cdot 10^3$	$-6.2 \cdot 10^3$
Environmental and biological quality	Climate change (kg CO2 eq.)	$-3.8 \cdot 10^4$	$-2.6 \cdot 10^5$	$-1.2 \cdot 10^5$	$-5.2 \cdot 10^5$	$-4.1 \cdot 10^5$	$9.6 \cdot 10^4$	$-2.9 \cdot 10^4$	$-4.5 \cdot 10^5$
	Nutrient balance (kg P eq.)	$-1.0 \cdot 10^1$	$-2.1 \cdot 10^2$	$-1.8 \cdot 10^2$	$-1.2 \cdot 10^2$	$-9.3 \cdot 10^1$	$-3.9 \cdot 10^2$	$-4.0 \cdot 10^2$	$-1.4 \cdot 10^2$
Human Health	Human toxicity (kg 1.4-DB eq.)	$-7.9 \cdot 10^3$	$-1.8 \cdot 10^5$	$8.8 \cdot 10^4$	$-9.7 \cdot 10^4$	$-7.8 \cdot 10^4$	$-3.0 \cdot 10^5$	$-3.2 \cdot 10^5$	$-2.5 \cdot 10^5$
Social welfare	Legality	1.1	1.1	0.7	1.0	2.7	0.3	0.7	1.0
	Perception -relevance	1.8	1.8	1.0	1.8	2.8	0.3	2.8	1.5
	Perception -influence on market	2.2	2.2	1.9	2.0	2.8	1.1	2.6	1.7

4.4 Discussion

Wicked problems – complex societal problems with multiple types of impacts to be considered, multiple perceptions on the problem and its optional solutions (Rittel and Webber, 1973) – ask for assessment approaches that transcend the classical risk assessment approaches. The latter often focuses on a single aspect of a problem, the reduction of which is used as target, while the transition towards a circular economy requires multi-metric sustainability assessments. Although sustainability is not a new concept and many sustainability assessment methods exist, there is room for improving the contribution of these assessments to decision-making (Benson, 2003; Kates et al., 2001; Little et al., 2016; Sala et al., 2013).

The protocol and the tool discussed in this manuscript facilitate a process that increases the value, transparency and reproducibility of sustainability assessments in decisions aimed at improved sustainability of alternative management strategies. Although there are often unknowns and the assessment can only provide the knowledge available to serve as one of the ingredients in decision-making (Harder et al. 2015), its contribution to the decision-making can be optimized with a process that is characterized by four elements:

1. Participation: the protocol makes it possible to select methods based on what is defined to be important by the different stakeholders, based on their expertise in disparate and often widely varying disciplines. This participation of stakeholders supports the scientific analyses of sustainability-relevant metrics by gathering all ideas of possible relevance and supports the decision process by gaining support for the assessment results.
2. Decision support: the method selection is discussed with the decisions that have to be made in mind, and with the stakeholders that play a role in these decisions.
3. Iteration: finding and deciding on solutions for wicked problems is characterized by adaptive management; the protocol (and the tool as well) can easily be used in an iterative way, transparently adjusting method selection and design based on new insights, e.g. from new stakeholders or knowledge.
4. Transparency: the protocol provides transparency in both the process of method selection as well as in what can be expected from an assessment (and what not).

The protocol and tool are flexible in design, which enables broader use. For example, aggregation across themes is not part of the protocol and the tool yet. If needed, aggregation of results from different sustainability metrics can be made by implementing value choices, which should in such cases be made transparent and explored in dialogue as well (Gasparatos and Scolobig, 2012; Özdemir et al., 2011).

The approaches that were proposed here have been applied in an example case study. The case study described a first iteration of the protocol with members of the ERF. The participants mentioned that the context evaluation provided a broad view on the situation, while the second phase, the question articulation, was structured such that it narrowed down to choices for themes and system boundaries, which were deemed necessary for method selection. The case study did not include participation of other stakeholders than the ERF members. However, the context evaluation resulted in the participants recognizing the importance to involve stakeholders in the process towards the SA and the interpretation of its results. This sequence, first with the internal organization, then an iteration with stakeholders, might be an effective working order in other case studies as well. Either for reasons of effort (as try-out) or to determine the point of view of the own organization before a discussion with external stakeholders is organized. External stakeholders identified during the context evaluation were the customers for the resources to be extracted, the government that decides on the juridical approval of the resource re-use and their use in new products and citizens, in order to check and act on citizen's perceptions of new products based on resources extracted from wastewater. The application of the protocol in the waste water case study revealed that the systematic approach towards context analysis and question articulation improved the quality and relevance of the items discussed on- and specified for the SA. As a result, next to the method selection, application of the protocol provided insight in the purpose of the assessments and in who are expected to act on the SA results in decision-making.

The tool, which was based on the review of existing SA methods, resulted in the identification of 30 currently available methods. This set can easily be supplemented with other methods. The scopes of the methods appeared to be unequal, which corroborates that method selection is a step that matters, i.e. it influences the assessment outcomes and thus decisions based on the outcome (Zijp et al., 2015). On the other hand, some of the methods have a comparable basis or overlap. For example, the Ecocost Value Ratio is in fact a combination of LCA, monetization methods and product costs. Some methods combine indicators with different scopes. For example, environmental impacts are covered at the whole life cycle of a product and social impacts only at the production stage. This is not necessarily wrong, but it should be transparently communicated when choosing to apply such methods. In the tool, this was operationalized with notifications during method selection. Regarding the sustainability themes, next to the well-known themes that fit within the areas of protection people, planet and prosperity, we found that technical issues are important to be additionally taken into consideration when deciding on alternative sustainable solution options and thus should be considered when designing a SA that supports these decisions. The review and the selection of methods in the tool was not extensive, but revealed the variability within the field of sustainability assessment methods for products. The tool was designed such that other methods can easily be added in the future, for example assessment methods that focus on other objects such as organizations. Also other system boundaries and themes may be added.

The case study showed that the protocol and the tool supported a thorough, transparent, case specific question articulation and method selection, which required iterations between 'demand' (the question) and 'supply' (the method and data availability). The selected methods were applied as first iteration in the process towards a strategic decision on which resources to focus on. The SA revealed, amongst others, that there is not one solution that scores best on all aspects that were defined to be important. Also, the solutions differed in their technical readiness level and thus in the information availability and with that the certainty of the derived scores. Hence, the decision for one or more of the solutions will have to involve weighing of different themes. The SA results showed that all solutions appeared to be beneficial for environmental quality and reduced resource depletion, compared to the business as usual scenario (Van Nieuwenhuijzen et al., 2016). Furthermore, one of the solutions seems unrealistic, given its chances in- and the identified perception of the market, and for this reason can be decided to be left out of further decision steps. In a next iteration, the stakeholders that were distinguished during the process can be invited to participate in the discussions on how to assess the different solutions. Also, the SA could be broadened by analyzing possible combinations of resource recovery (De Fooij, 2015; Van Nieuwenhuijzen et al., 2016) and taking into account actual yield estimates at the different WWTP's in the Netherlands (Van Nieuwenhuijzen et al., 2016).

The application resulted in three general lessons for method selection. First, there was not one existing method that fitted all the requirements that were made explicit by the ERF. Hence, a combination of methods was required for the assessment. This will often be the case. Secondly, method selection appeared to require an iterative process. In the case study, the method selected was not fully applicable due to limited data availability and therefore, after consideration between the SA experts that performed the assessment and the ERF, that part of the method was replaced by another approach. Thus, after a method is selected that fits the question articulation, other elements of method selection, e.g. data availability, software availability and costs of the assessment, need be explored in order to check if the method can be applied. This can lead to adjustment of the method (this case study) or the selection of another method. Finally, the application of the protocol and tool presented in this manuscript revealed the important role of scientists and sustainability experts. Their role is larger than the technical role of performing the sustainability assessment. Firstly, based on knowledge and experience in the field of sustainable development they can contribute in the discussion on what is important to include, as one of the stakeholders. They have to guard that the method selection is not arbitrary, but aligned with the state of the art knowledge with regard to sustainable development. Secondly, the specification can lead to a list of selected sustainability themes that partly overlap. The expert can propose a coherent set of themes based on the input of the question articulation. Thirdly, when the list of possible (combinations of) methods are derived the expert can propose which combination suits best, given the organizational restrictions such as data availability. Finally, experts can take the initiative at every opportunity by questioning the question and apply, evaluate and improve the approaches provided in this manuscript.

4.5 Conclusions

Existing efforts towards designing and implementing circular economy principles demonstrate the urgent needs to expand mono-univariate risk prevention approaches aimed at risk prevention and reduction to multi-metric sustainable-development approaches aimed at swift sustainability improvements, both in science as well as in (decision-support) practice (Zijp et al., 2016). This need is not exclusive for the transition towards a circular economy, but extends to all sustainability development goals (UN, 2015a) and to all the wicked environmental problem definitions decision makers at every policy level are confronted with. The field of sustainability science is developing swiftly, but requires improvement on the assessment process and its contents: i.e. question articulation, linking metrics to the question at hand, facilitating stakeholder participation and finally providing decision support under (multi-metric) outcomes and associated uncertainty. The protocol and tool presented here provide a practical way to manage and execute the process of selecting a SA method with input from stakeholders. Based on the case study, we conclude that the proposed approaches can support users in managing the processes of question articulation and method selection. They are an operationalization of the recently proposed SA identification key (Zijp et al., 2015). Both the protocol and tool have a flexible design, which enables broader use and further development based on growing experience and new insights. Using them increases the chance that SA outcomes are used in the decision-making context and indeed contribute to reaching goals for sustainable development at all levels.

4.6 Acknowledgements

This work was funded by RIVM's strategic research program, run under the auspices of the Director-General of RIVM and RIVM's Scientific Advisory Board, under the projects Measurably Sustainable S/607020 and SafeBBE S/124001. We thank the members of the ERF that contributed to the workshops, the experts that contributed to the expert elicitation and Jeroen Devilee and Mark Huijbregts for their comments on draft versions of this manuscript.

4.7 Annexes

4.7.1 Terminology

Scientific literature describing sustainability assessment methods use a plethora of terms, such as instrument, tool and method; or theme and indicator. Terms seem interchangeable, but they can have different meanings in different articles, which can cause confusion. Therefore, the definitions used in this paper for some of those words are listed below:

- Themes: aspects that are thought to be considered in sustainable decision-making, e.g. contribution to climate change, water depletion or slavery.
- Capital: areas of protection. Capitals are used to organize the themes under the capitals they contribute to (Fig. 3 in the manuscript).
- Indicator: a parameter or a value derived from parameters, which point to, provides information about, or describes the state of a theme, with a significance extending beyond that directly associated with its value (OECD, 2003).
- Method: a collection of consecutive and/or complementary sub-methods with which a specific question can be answered. Examples of methods are Life Cycle Analysis (LCA), Ecological Footprint (EF) or EMergy analysis (EM).
- Sub-method: sub-methods are the consecutive and/or complementary analytical steps a method consists of. For example methods with which indicators can be quantified or with which results can be aggregated (Zijp et al., 2015).
- Domain: domains are the organizing structure of an identification key. Criteria that characterize the object of the identification key, in our case sustainability assessments, are grouped into domains (Nickerson et al., 2013).

4.7.2 Criteria for method selection

Table S4.1 Criteria for sustainability assessment method selection, collated from literature by Zijp et al. (2015) and used in this study to analyze the inventory of sustainability assessment methods. The first two columns are taken from Zijp et al. (2015). The third column indicates whether and where the criteria are used in the method selection tool and the method selection protocol derived in this study. Aggregation out various output metrics into e.g. a comprehensive score was not in the scope of this study and can be added later to the tool and the protocol.

Criteria	Explanation	Covered by tool and/or protocol
Domain: System boundaries/Inventory		
Object	What is the object of the assessment? *Is it a physical object (product, chemical, process), or an organization, a* region, a policy measure, an activity, etc…	Tool and Protocol
Spatial focus	What is the spatial focus of the activity? *Is the activity assessed on the local scale (site-specific data), regional scale (region-generic data), worldwide or anywhere (generic data), or a combination (partly location-specific and partly worldwide or anywhere)?*	Tool and Protocol
Temporal focus	What is the temporal focus of the assessment? *Is the activity assessed retrospectively, prospectively or does a snapshot suffice?*	Tool and Protocol
Life cycle thinking	Which parts of the life cycle or supply chain are included in the assessment? *No life cycle, the whole life cycle (cradle to grave), or something in between (cradle to gate)?*	Tool and Protocol
Domain: Impact assessment/ Theme selection		
What is to be sustained	What are the areas of protection? *Economic welfare, environmental and biological quality, resources, social welfare, human health and/or technical welfare?*	Tool and Protocol
Theme and indicator selection	Which themes are selected? (see Fig. 3)	Tool and Protocol
Spatial focus of impact	*What is the spatial scale of the impacts that should be taken into account?* Does the assessment include intra-generational impacts? Or in other words: does the assessment aim at internal or external sustainability. Impacts at what scale are taken into account? Are they site-specific/dependent or independent?	None
Temporal focus of the impact	*What is the temporal scale of the impacts that should be taken into account?* Does the assessment include inter-generational impacts? What time-frame should be included for the impacts?	None
Domain: Aggregation/Interpretation		
Out of scope	Out of scope	None
Domain: Method Design		
View on stakeholder involvement	Who should be involved in the assessment in which way? Also referred to as legitimacy, in relation to indices or composite indicators.	Protocol
Context of the assessment	How and by whom are the results used? In which (phase of a) procedure are the results of the assessment used? Is the goal of the measure: decision-making and management, advocacy, participation and consensus building or research and analysis? Or is it a strategic, capital investment, design and development, communication and marketing, or operational question?	Protocol
Uncertainties	How are uncertainties to be handled? Salience, credibility and variability? Should an uncertainty, sensitivity and/or perturbation analysis be included?	None

Table S.4.2 Definition of the six areas of protections, or so-called capitals, used for the method inventory and the tool, with example-description of their meaning. The capitals are based on subjects of existing methodologies, which are referred to under 'refs' (literature references), the number corresponds to the sources listed in Table S4.5. Capitals may be part of more methods than listed in this table, for an overview see Figure S4.1 (Annex §4.7.5).

Capital	Description	refs
Economic welfare	What is the impact on the various economies involved? Includes the availability of work, etc.	19, 27 & 32
Environmental and biological quality	How is the environment, both biotic as well as abiotic, affected due to the activity? E.g. climate change, soil erosion etc.	2, 13, 15, 18, 23
Human health	What impacts on human health are associated with the activity.	3, 15, 23
Social welfare	What are the rights for the people involved in the production chain.	13, 26, 27, 33
Resources	How much land, water, energy and materials are needed for the activity?	2, 5, 6, 18
Technological welfare	Technical issues that are required in order to make a new process or product sustainable are for example flexibility (can the process design be adapted, e.g. when substances are not available anymore) and feasibility (is a full-scale operation feasible, with regard to availability of required feedstock etc.).	1, 28, 30,33

4.7.3 Protocol for workshops

Table S4.4 Questions for the question articulation phase. Question 2 to 7 are also asked in the online tool and used to select a sustainability assessment method.

No.	Question
1	What is the goal of the assessment? Where should the assessment contribute to?
2	What is the central object of the question (a product, product group, organization, …)?
3	What is the technical readiness level of the object (lab, development, pilot, commercial)?
4	What is the spatial scale the question focusses on (local, regional, national, global)?
5	What is the temporal scale the question focusses on (present, past, close future, far future)?
6	Which parts of the life cycle should be included?
7	Which sustainability themes are important to include in the assessment?
8	Which stakeholders that are not yet involved should be involved in this case study?
9	Will they have another view on the questions above then you just answered?

Table S4.3 Protocol for guiding the process of precise question articulation, given a sustainability-oriented problem definition (opportunity).

Part	Goals	Example activity
Evaluation of context and role of sustainability within the context	Get insight in the context of the assessment. Participants get insight in how each of the stakeholders views the situation. Participants get insight in the context of the sustainability assessment (it is not something on its own). Participants get insight in their own- and the other's views on sustainability.	Introduce yourself and use an object you identify with sustainability. Find one or two persons you do not know very well and interview each other on what makes proud or angry concerning the situation. Brainstorm individual on what you would like to see solved within your organization. Write all individual results on separate papers and cluster them together. Discuss which of the clusters are sustainability questions and why.
Question articulation	Specify the question and define sustainability such that a (set of) methods can be defined that matches the question.	Go into groups of 6-8 participants, with one process manager and someone that makes notes. Drop the central question and try to answer the underlying types of questions in Table S4.4 together. Try give one concrete answer. When there is an argument on what should be the answer score both answers as appropriate. When all groups are finished present each other the results and discuss differences. Try to harmonize the results. When this seems not possible leave the different options open.
Search for a (combination of) pertinent method(s)	Select a (set of) method(s) and refine it. Make an action plan on the route towards a decision.	Use the outcomes of the question articulation to fill in the tool. Discuss the resulting methods (an expert can present the different methods). Do the different options left open in the previous round result in different methods or not? Is there one method that fits all requirements or is a combination required? Make a timeline: when are the assessment results necessary and who are expected to use the results. Plan back who should be involved then and when the different phases of an SA should take place (goal/scope definition, inventory, impact assessment, interpretation etc.) and who should be involved in the different phases. Use also the questions and answers of Table S4.4, i.e. question 1 and 8.
Optionally: Refine method selection	Check if selection can be executed and based on that refine method selection.	Of the selected (combination of) method(s) find out what the data, expertise and model (software) requirements are and, if necessary, adjust the design according to your findings, in accordance with the participants of the question articulation phase.

4.7.4 Inventory of methods and extra references used for categorization of themes

Table S4.5 Inventory of currently existing sustainability assessment methods (1-31, alphabetically sorted) and extra references used for categorization of themes (32-33)

	Name	Source (accessed at 03-11-2016)
1	Bioref-Integ	http://iturl.nl/snabLIU
2	Cramer Criteria	http://iturl.nl/snd0cdS
3	Decision framework for chemical process design*	http://iturl.nl/snj--gW
4	Driving Force Analysis**	http://iturl.nl/sn7iWar
5	Eco-Costs / Value Ratio (EVR)	http://iturl.nl/sn1Z8Hq
6	EU Ecolabel	http://iturl.nl/snIRISA/; http://iturl.nl/snEofuY
7	Ecological Footprint	http://iturl.nl/snZHwTb
8	Ecoscale	http://iturl.nl/snoj-HH
9	Ecover Diamond Model (method not disclosed)	http://iturl.nl/snKhl91
10	E-factor***	http://iturl.nl/sn9t06N
11	Emergy analysis	http://iturl.nl/sneiC5h
12	Exergy analysis	http://iturl.nl/snbxD16
13	Fairtrade Standards for Small Producer Organizations	http://iturl.nl/snQtW
14	Global material economy****	http://iturl.nl/snh2DaW
15	GreenScreen for Safer Chemicals	http://iturl.nl/snIhe
16	GreenWERCS (not free)	http://iturl.nl/snRIEHD
17	Guide on Sustainable Chemicals	http://iturl.nl/snIc6J4
18	Life Cycle Assessment (LCA) with ReCiPe	http://iturl.nl/snj-rY7
19	Life Cycle Costing (LCC)	http://iturl.nl/snqMXg4
20	Life Cycle Metrics for Chemical Products	http://iturl.nl/snW95
21	Material Flow Analysis	http://iturl.nl/snVNBJB
22	Material Input per Service Unit (MIPS)	http://iturl.nl/snn9g
23	PBT Profiler	http://iturl.nl/snsHiE-
24	RSB Principles & Criteria for Sustainable Biofuel Production	http://iturl.nl/snYYdg7
25	Product Environmental Footprint (PEF)	http://iturl.nl/sndAFXq
26	Prosuite	http://iturl.nl/sn2iy website offline
27	Social Life Cycle Assessment (SLCA)	http://iturl.nl/sntD4xl http://iturl.nl/snaTsLb
28	Sustainability Assessment of Technologies (SAT)	http://bit.ly/2qunW3P
29	Sustainability metrics: LCA and Green Design*****	http://bit.ly/2qHtPHo
30	GBEP Sustainability Indicators for Bioenergy	http://bit.ly/2pI2DZL
31	The Higg Index	http://bit.ly/2pACk9z
32	Global reporting initiative	http://bit.ly/1aVFfc3
33	Dow Jones Sustainability Index	http://bit.ly/1uprsXi

* Decision framework for chemical process design: Sugiyama et al 2008 AIChE, 54, 1037-1053, scientific literature, not free
** Driving Force Analysis: Obenndip and Sharratt 2006 Org Process Res Dev, 10, 430-440, scientific literature, not free
*** The E Factor: fifteen years on: Sheldon 2007 Green Chem, 9, 1273-1283, scientific literature, not free
****Global material economy: Augé and Scherrmann 2012 New J Chem, 36, 1091-1098, scientific literature, not free
***** Sustainability metrics: LCA and Green Design: Tabone et al. 2010, Environ Sci Technol, 44, 8264-8269, scientific literature, not free

4.7.5 Coverage of Capitals and themes by thirty existing sustainability assessment methods

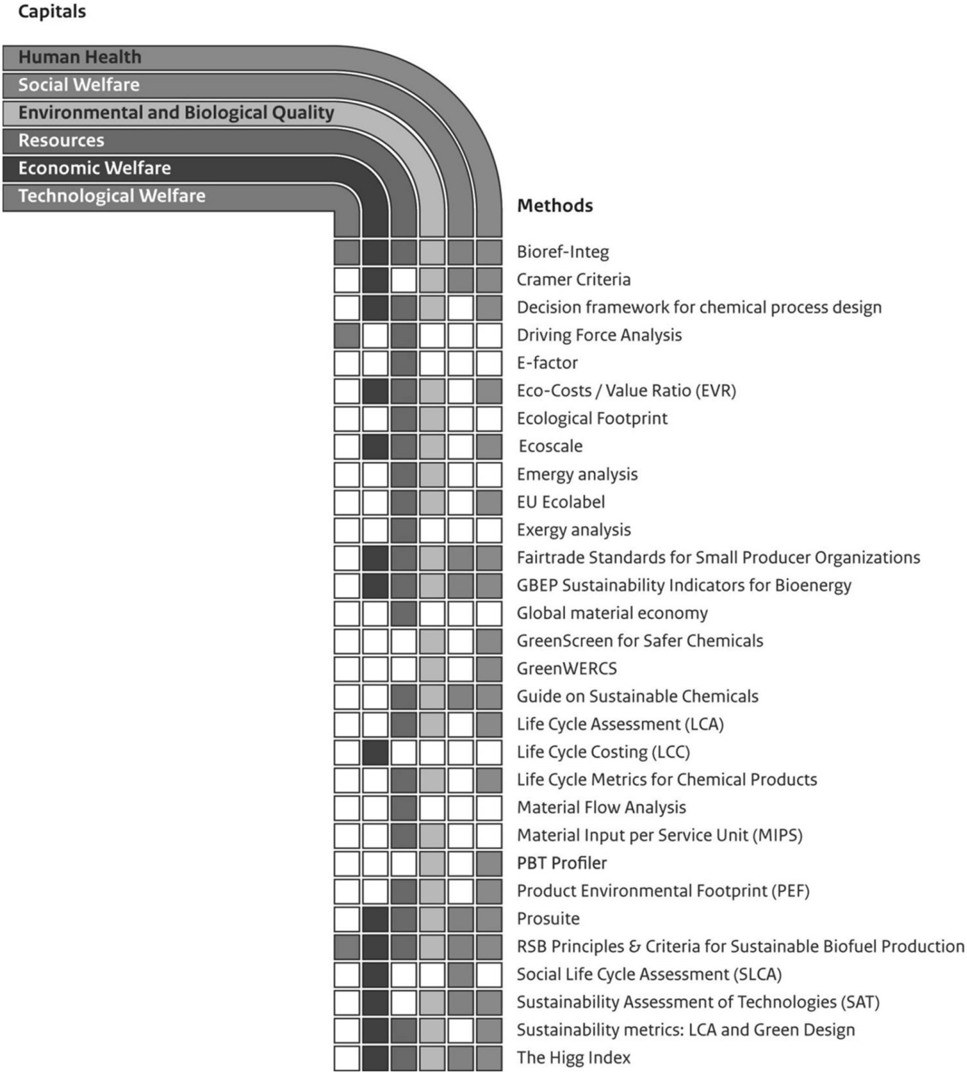


Fig. S4.1 Coverage of capitals (areas of protection) per method. Please note that the requirements for the EU Ecolabel and Fairtrade Standards are product category or sector specific. Therefore, sustainability assessments of specific products or sectors may include additional or adapted areas of protection (capitals or themes) for the Ecolabel or Fairtrade Standard.

4.7.6 Conceptual model of a generic WWTP in the Netherlands detailing process steps at which

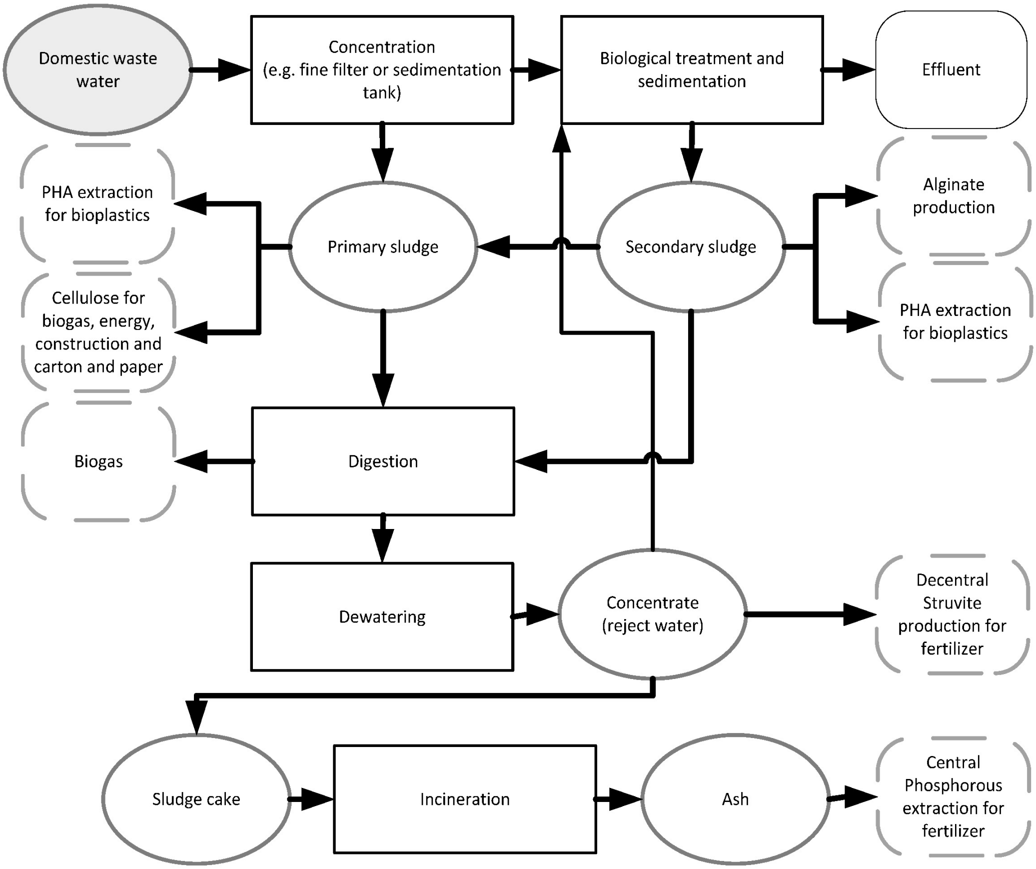


Fig. S4.2 Conceptual model of the resources that can be recovered from waste water in different steps of a general waste water treatment plant in the Netherlands (based on STOWA (2016); Van Nieuwenhuijzen et al. (2016)). Rectangles represent processes in the domestic waste water treatment, ovals are different stadia of the waste (water) flow, and rounded rectangles represent recovered resources and their production process. The selection of either of the possible end-products can limit the resource recovery on alternative pathways.

4.7.7 Literature used on economic costs and benefits of the solutions

1. Becker N. Installation of SRS (Sewage Recycling System) at Aarle-Rixtel WWTP. A joint report with Waterschap Aa en Maas (WSAM). November 2014 (http://edepot.wur.nl/326365)
2. Bixler HJ, Porse H. A decade of change in the seaweed hydrocolloids industry. J Appl Phycol, 23(3), 321–35, 2011.
3. Bluemink E, van Nieuwenhuijzen A, Wypkema E, van der Kooij Y. Bioplasticproductie uit zuiveringsslib vereist marktgedreven procesoptimalisatie voor economisch rendement. H2O-Online [Internet]. mei 2014; Beschikbaar op: http://bit.ly/2plaRBf
4. Dewaele C. NuReSys. From P recovery to fertilizer production. 2015 mrt; 2nd European Sustainable Phosphorus Conference (ESPC2).
5. E.D. Bluemink, A.F. van Nieuwenhuijzen, E. Wypkema, C.A. Uijterlinde. Bio-plastic (Poly-hydroxy-alkanoate) Production from Municipal Sewage Sludge, a Technology Push or a Demand Driven Process? Water Sci Technol, 74(2), 353-358, 2016
6. Gudde T. Notitie aan Energie en Grondstoffen Fabriek, kenmerk WO-1410-0131-mzf. 2014.
7. Molinos-Senante M, Hernández-Sancho F, Sala-Garrido R, Garrido-Baserba M. Economic Feasibility Study for Phosphorus Recovery Processes. AMBIO, 40(4), 408–16, 2011.
8. Nieminen J. Phosphorus recovery and recycling from municipal wastewater sludge [Internet] [Master thesis]. [Espoo]: Aalto University; 2010 [geciteerd 20 juli 2016]. Beschikbaar op: http://bit.ly/2pbwzfE
9. Petzet S, Cornel P. Prevention of Struvite Scaling in Digesters in Combination with Phosphorus Removal and Recovery - The FIX-Phos Process -. Technische Universität Darmstadt; IWAR;
10. Petzet S, Peplinski B, Bodkhe SY, Cornel P. Recovery of phosphorus and aluminium from sewage sludge ash by a new wet chemical elution process (SESAL-Phos-recovery process). Water Sci Technol, 64(3):693–9, 2011.
11. STOWA 2012-07. Verkenning naar mogelijkheden voor verwaarding van zeefgoed. Amersfoort; 2012.
12. STOWA 2013-32. Fosforhoudende producten uit de communale afvalwaterketen wet- en regelgeving, marktkansen, verwerkingsconcepten. Amersfoort: STOWA; 2013.
13. STOWA 2014-10. Bioplastic uit slib verkenning naar PHA-productie uit zuiveringsslib. Amersfoort: STOWA; 2014.
14. STOWA 2016-12. Marktverkenning en gewasonderzoek struviet en struviethoudende producten uit communaal afvalwater. Amersfoort: STOWA; 2016.
15. STOWA 2016-22. Levenscyclusanalyse grondstoffenfabriek. Producten uit de RWZI STOWA, RvO, Grondstoffenfabriek; STOWA report 2016- 22, ISBN 978.90.5773.713.8, 2016
16. Tchobanoglous G, Abu-Orf M, Metcalf & Eddy, Inc, AECOM, redacteuren. Wastewater engineering: treatment and resource recovery. 5. ed. New York, NY: McGraw-Hill; 2014. 2018 p.
17. Van den Heuvel G. Nereda zuiveringstechnologie verovert de wereld. Nemo Kennislink. mei 2012;

4.7.8 Details of the perception analyses

The following procedure was followed for the expert elicitation on the influence of perception and legality on the chances of the business case for resources from waste water. The procedure is based on the description in the Prosuite Handbook (Gaasbeek and Meijer, 2013).

Select the experts

Criteria:

- Include at least three individuals with academic or practical experience in the field
- When possible include persons from different disciplines and different professions (i.e. research, consultancy, government, NGO's, industry). This should minimize unintended biases
- Do not attempt to exclude experts with contradictory viewpoints

Results:

For every option three or more experts contributed to the elicitation: at least one decision maker from the board of the water boards, one market expert and one expert that was involved in the pilots for the resource recovery.

Document results

Criteria:

- Provide complete and explicit results of the consultation and
- Summarize the overall conclusion
- Identify and recognize uncertainties and document them.
- Use performance reference points with clear definition

Results:

Table S4.5 lists the questionnaire and the performance reference point system.

Overall conclusions are mentioned in the main text of the manuscript. The final scores were derived by averaging the scores per expert.

Concerning the uncertainty: given the premature phase of the scenarios some of the questions were hard to answer by the experts. In those cases, the answers were generated from general experience of respondents with other pilots than with case-specific knowledge. Nonetheless, the experts appeared to provide relatively uniform scores. Furthermore, the research was within the minimum requirements but a larger set of respondents could provide a different picture. E.g., questioning civilians and potential producers would be interesting. It is noted, however, that some of the experts that did participate have years of experience in the field, including contact with potential users, evaluations of perceptions and of legal aspects as potential limiting factor. Additional lessons were drawn from the qualitative information provided in the process. Still, the quantification allows comparison between the options: there is a clear difference between the options in the way perception and legality influences a methods' potential for implementation.

Table S4.6 Questions asked and scoring of possible answers for the perception analyses. During the interviews, there were also open questions of which the answers are summarized as qualitative results.

	Question	Answer	Score
1	How important is perception of safety for using the resource, compared to other issues like sustainability, costs, continuity in availability and quality of the resource	Not important	3
		Important, but less than other issues	2
		Important, equal to other issues	1
		Important, more than other issues	0
2	What is the influence of the perception of producers that potentially use the resource in their product, on the chances of the resource on the market	Positive	3
		No influence	2
		Threshold that can easily be over won	1
		Negative	0
3	For what part of the potential producers is this influence negative?	Less than 25%	3
		Between 25 and 50%	2
		Between 50 and 75%	1
		More than 75%	0
4	Will the perception of the producers expectedly change in the positive direction?	Yes on short term (<5 yrs)	3
		Yes on long term (>5yrs)	2
		No	1
		No, in the opposite direction	0
5,6,7	*Question 2-4 but then for consumers instead of producers*		
8	What is the influence of current regulations on the chances of the use of a resource from waste water compared to their alternatives?	Positive influence	3
		No influence	2
		Negative influence that will be over won	1
		Negative influence	0

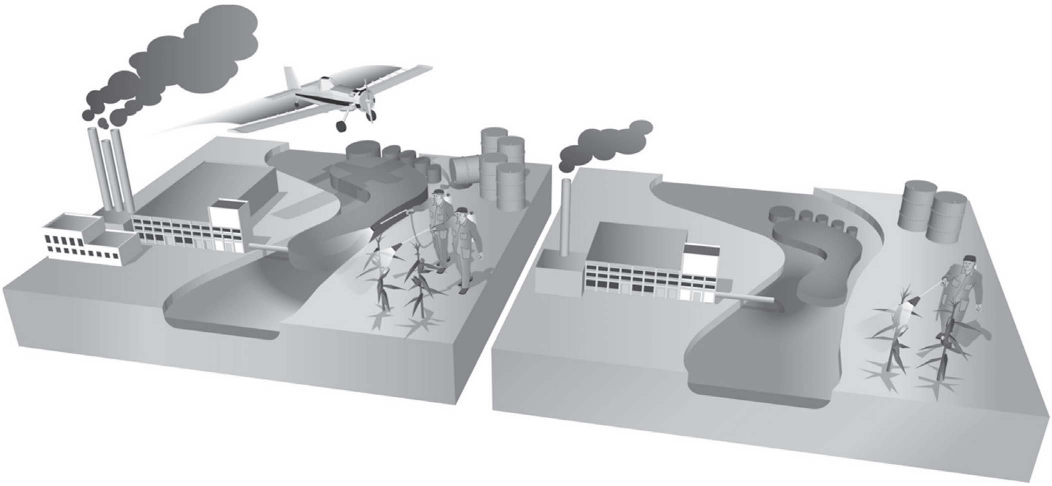
Qualitative results

- For most options safety was mentioned as one of the least concerns compared to other issues, like costs (has to be equal or lower compared to alternative) and quantity and quality (is there a constant resource flow with sufficiently constant quality).
- Often, especially the perception of companies that use the resource was mentioned as important. When user prescriptions are clear and the product 'looks and feels' good, most consumers are expected to trust the system (the producer and government) and do not worry about the safety, provided that neither calamities or negative news happen.
- When a calamity happens this was expected to seriously change (harm) the perception and with that the chances on a business case.
- Legality was mentioned as serious problem for market penetration of novel approaches. To overcome this, the waste label currently associated to the resources (in not only the Netherlands but also the rest of Europe) was suggested to be changed into a product label.
- Discussions on potential residues of medicines in the products are viewed as problematic for beneficial perception and regulation. This was suggested to require good research and communication.
- All experts were optimistic about the change in perception in a positive direction (allowing resource utilization), except for the use of cellulose from waste water in the paper and carton industry. Most, but not all experts mentioned that they expected no market adoption of that option.
- The influence of perception and legality was especially earmarked as a problematic issue if the resource will be used in the food chain.
- Replies implied that water boards should work together, in order to be able to guarantee sufficient resource continuity.
- It was mentioned that it is not necessary that everybody agrees on using products from waste water. Only a small market was deemed sufficient to make a resource utilization strategy successful.

5 Definition and applications of a versatile chemical pollution footprint method

Authors: Michiel C. Zijp, Leo Posthuma, Dik van de Meent

Published in: Environmental Science and Technology (2014) 48, 10588–10597



Abstract

Due to the great variety in behavior and modes of action of chemicals, impact assessment of multiple substances is complex, as is the communication of its results. Given calls for cumulative impact assessments, we developed a method that is aimed at expressing the expected cumulative impacts of mixtures of chemicals on aquatic ecosystems for a region and subsequently allows to present these results as a chemical pollution footprint, in short: a chemical footprint. Setting and using a boundary for chemical pollution is part of the method. Two case studies were executed to test and illustrate the method. The first case illustrates that the production and use of organic substances in Europe, judged with the European water volume, stays within the currently set policy boundaries for chemical pollution. The second case shows that the use of pesticides in Northwestern Europe, judged with the regional water volume, has exceeded the set boundaries, whilst showing a declining trend over time. The impact of mixtures of substances in the environment could be expressed as a chemical footprint, and the relative contribution of substances to that footprint could be evaluated. These features are a novel type of information to support risk management, by helping prioritization of management amongst chemicals and environmental compartments.

5.1 Introduction

Approaches to quantify cumulative impacts of chemical exposures are needed for useful prioritization of environmental policy and management efforts and ample room for major improvements in this field regarding both science and communication has been shown (Bu et al., 2013; Giubilato et al., 2014). This paper describes a basic sequence of steps to derive cumulative environmental impacts of chemical pollution and expresses the output as a 'polluted volume' and as a pollution index. Those results can be interpreted and communicated as 'chemical pollution footprint' (in short: chemical footprint). The four main triggers to explore the possibility and applicability of expressing the environmental impact of multiple substances in a footprint measure were: i) latitude to improve cumulative impact assessment, ii) the definition of 'holistic' policy goals, like e.g. the good ecological status in the EU Water Framework Directive(EC, 2000), which require information on the impacts of mixtures of substances in addition to the traditional per-substance knowledge(EEA, 2011); iii) evidence for large variability in current exposures in different environmental compartments (USEPA, 2009); and iv) discussions in literature on chemical pollution as one out of nine systems for which planetary or regional boundaries defining the safe operating space for humanity should not be transgressed (Rockström et al., 2009a), and the linked notions on the complexity of predicting the impact of mixtures of chemicals (Rockström et al., 2009b) in relation to the definition of a boundary on biodiversity loss (Persson et al., 2013). Furthermore, it is suggested that a chemical footprint method should be added to the so-called footprint family(Fang et al., 2014) and that the combination of a carbon footprint, ecological footprint and toxicity impacts (chemical footprint) gives a relative complete indication of the environmental part of sustainability (Mattila, 2013).

This study encompasses a definition of - and a stepwise approach to derive cumulative impacts of chemical pollution in an area, expressed as a chemical footprint for biodiversity impacts. This fills a current gap, is in line with other footprint methodologies, and it addresses criticisms on footprint methodologies, e.g. concerning the ecological relevance of the system boundaries of the assessment and of the weighting procedure(Van den Bergh and Grazi, 2013). The method is versatile, and can be used for other scales of emissions and endpoints too, since it essentially checks whether there is sufficient 'dilution volume' in the environment to remain safe.

Footprint methodologies exist for various impacts and pressures (Čuček et al., 2012). Novel concepts and applications are published in both scientific literature(Gleeson et al., 2012) and in policy documents (EC, 2013). The general concept of footprinting is used for the evaluation of various impacts of human activities and encompasses a strong communicative element. Since the introduction of the term 'ecological footprint' in 1992 (Rees, 1992), footprint methodologies on more than 26 environmental issues have been developed, of which the ecological footprint, the water footprint, the carbon footprint and the energy footprint are the most frequently used (Fang et al., 2014). In general a footprint has been defined as (Čuček et al., 2012): "A quantitative measure describing the appropriation of natural resources by humans". Appropriation refers to the central principle of footprints that all natural resources necessary to support the human activities under consideration should be included. Specified for the chemical footprint we propose the following definition (based on Environmental Space Theory (Adriaanse, 1993) and the Grey Water Footprint (Hoekstra et al., 2009)):

"A quantitative measure describing the environmental space needed to dilute chemical pollution due to human activities to a level below a specified boundary condition".

Conceptually, such boundary conditions can be related to classically used environmental quality standards as applied in preventive chemical regulations(EC, 2006), or to impacts observed or estimated in ecosystems, e.g. tipping points (Scheffer et al., 2009).

Chemical footprinting has been discussed in literature for a while (Hoekstra et al., 2009; Panko and Hitchcock, 2011; Sala and Goralczyk, 2013). Apart from single-chemical footprints such as proposed by Hitchcock et a l(Hitchcock et al., 2012; Panko and Hitchcock, 2011) and pressure-focused chemical footprints such as the Grey Water Footprint (Hoekstra et al., 2009), the chemical footprint of impacts due to cumulative emissions to an area is novel still undefined, especially regarding the boundary condition(Sala and Goralczyk, 2013). By combining earlier works in cumulative impact analyses (Harbers et al., 2006; Henning-De Jong et al., 2008) with ecological analyses and boundary proposals, we now propose a comprehensive versatile method to derive a chemical footprint.

The method proposed here consists of classical elements of chemical risk assessment (emission analysis, single compound exposure assessments and effects assessments), expanded with mixture impact assessment and quantitative approaches for defining the boundary condition and 'environmental space occupied'. Mixture toxic pressure was quantified as multi-substance Potential Affected Fraction of species (msPAF) (De Zwart and Posthuma, 2005; Posthuma et al., 2002), which closely relates to biodiversity impacts (Posthuma and De Zwart, 2006; Posthuma and de Zwart, 2012; Smetanová et al., 2014). Others already used this method to predict ecological impacts of 343 high production volume chemicals and of the use of pesticides in the combined catchments of the rivers Rhine Meuse and Scheldt (RMS) the RMS catchment (Harbers et al., 2006; Henning-De Jong et al., 2008).

Building forth on this work, and in line with general footprint definitions and considerations on protective chemical policies and boundaries for the safe operating space of humanity, the aims of this study are i) to propose a versatile method to express the expected ecotoxicological impact of a mixture of substances in the environment, and express the results as a chemical footprint; ii) to apply the method in two case studies; and iii) to discuss the results and explore its usability for policy.

5.2 Methods

5.2.1. Calculation procedure.

The aforementioned definition of a chemical footprint encompasses the elements of (1) exposure assessment, (2) impact assessment, (3) the boundary condition, and (4) the dilution volume needed to reach the set boundary condition. Given these elements, the developed chemical footprint method consists of the following procedure:

1. Target and scope: define compounds and areas of interest;
2. Quantification of emissions: quantify emissions to the environment (unit: g);
3. Quantification of the aquatic exposure to chemical pollution (unit: g/L);
4. Quantification of the ecological impact due to chemical pollution (unit: % of species affected);
5. Quantification of chemical pollution boundary (unit: % of species affected);
6. Express the results as Footprint (unit: m3), or as an index, relative to a selected available environmental space (unit-less).

The technical design of the procedure's individual steps comprises both existing methodologies and new elements, as detailed below.

1. Target and scope. The target and the scope of the assessment must be chosen. This determines the conditions for the following steps, and it is key to the final interpretation. The following four types of questions delineate the assessment and the footprint interpretation generated:

I. Which human activity is under consideration? Is the object e.g. a chemical, a product, a region or a policy measure
II. What question is to be answered? Is the required outcome absolute, e.g.: what is the impact of an activity? Or is it relative, e.g.: which chemicals/ activities/ sectors/ etc. contribute most to the total impact? Or: How does the impact change over time, given altered use patterns, related to policy measures planned or taken.
III. What are the system boundaries of the activity? For example: which phases of the chemicals life cycle are included (raw material extraction, production, use, disposal)? Or: what is the spatial scale

of the activity? Is it local, regional, world-wide, or organizational; and what is the temporal scale of the activity e.g. seasonal and yearly; and

IV. What are the receptor endpoints (e.g., biodiversity), and on which scale are the endpoint impacts considered? Does the assessment address ecological integrity in water, soil, sediment or air, and on a local, regional, or global scale?

2. Quantification of emissions. Next, the amounts of chemical substances, released into the environment as a result of the activity considered, within the set scope of the assessment, need to be quantified (E in g yr^{-1}). To focus the assessment and limit data gaps it could be considered to include only a smart selection of the original list of substances in the assessment, for example only the chemicals with the highest volumes (Harbers et al., 2006) and/or most hazardous characteristics (Strempel et al., 2012). In practice, emissions often need to be estimated by application of use- and substance-specific emission factors (EF)(CEFIC, 2012), because measurements are often not available for all the chemicals and geographical scales relevant for the assessment. Production volumes may be available, although not always accessible (confidential), e.g. from REACH dossiers.

3. Quantification of the exposure. The concentrations of the released substances in the different receptor compartments on different geographical scales ([C] in g l^{-1}) need to be estimated. Multimedia fate models were developed and are applied for this purpose since 1978 (MacLeod et al., 2010). These models generate steady state concentrations in various environmental compartments, given compartment and compound characteristics. The results of the model can be validated with field data. When these are not available, transparent reporting of model choices, input and output, accompanied with uncertainty and sensitivity analysis becomes essential according to Good Modeling Practice Guidelines proposed by Buser et al. (2012). In the case studies, we built forth on earlier studies (Harbers et al., 2006; Henning-De Jong et al., 2008) without major changes (Table 5.1). We used the nested multimedia mass balance model SimpleBox (Den Hollander et al., 2004) to estimate compartment concentrations.

4. Quantification of ecological impact. Mixture toxic pressure was derived from the concentration data using the mixture models developed by De Zwart and Posthuma (2005), using EC50 data (acute), because EC50s have been shown to represent actual damage in aquatic ecosystems (Posthuma and De Zwart, 2006; Posthuma and de Zwart, 2012; Smetanová et al., 2014). Further details are described elsewhere (De Zwart and Posthuma, 2005; Harbers et al., 2006; Posthuma et al., 2002). Differences with earlier studies (Harbers et al., 2006; Henning-De Jong et al., 2008) are summarized in Table 5.1. Major differences are the use of lognormal instead of log-logistic SSDs and the calculation of the uncertainty of species variance in toxicity. Lognormal SSDs were preferred above the log-logistic distribution functions because of reasons summed by Van Zelm et al. (2007): abundant use in many scientific fields, the central limit theorem and the phenomenon that the use of a log-logistic distribution function can result in unrealistic effect factor values. The uncertainty of species variance in toxicity was calculated per Toxic Mode of Action (TMoA) like in Henning-de Jong et al., while Harbers et al., assumed a dispersion factor of 2 for each TMoA.

Table 5.1 Comparison of methodological choices by Harbers et al., Henning-de Jong et al. and this study.

Subject	Harbers et al.	Henning-de Jong et al.	Case study 1 and 2
PAF calculation	Log logistic	Log logistic	Lognormal
Toxicity data	EC50, acute	EC50, acute	EC50, acute
Uncertainty of species variance in toxicity	Based on dispersion factor of 2 for each TMoA	Derived for each TMoA individually	Derived for each TMoA individually
Receptor	Regional marine water	Regional marine water	Fresh and marine; regional, continental and global
Use of prior data in uncertainty analysis*	No	No	Yes

*Details below, under 'uncertainty assessment', paragraph 5.2.4, and in the SI at: http://pubs.acs.org/doi/suppl/10.1021/es500629f

5. Quantification of chemical pollution boundaries. Adriaanse (1993) mentioned two elements needed to evaluate 'appropriation of natural resources': 1) a sustainability level or policy target expressing which impact is still acceptable (also referred to as a boundary for chemical pollution (Posthuma et al., 2014; Rockström et al., 2009a), and as 'carrying capacity' by Sala and Goralczyk (2013), and 2) the amount of "environmental space" of the system under consideration.

A boundary for chemical pollution is ideally based on natural thresholds that express the vulnerability of an ecosystem (Brook et al., 2013; Persson et al., 2013; Posthuma et al., 2014; Rockström et al., 2009b). The safe boundary must be below such a threshold, in order to account for uncertainties and other stressors, to eventually prevent impacts. As such, the boundary concept encompasses both science (thresholds) and value judgment. In our case the boundary should be expressed as the fraction of species above which ecosystems are expected to be affected by chemicals to an irreversible extent (msPAFmax). The use of such a boundary for footprint exercises has recently been mentioned by Fang et al. (2014), Sala and Goralczyk (2013) and Posthuma et al. (2014) but has not yet been applied. Examples of ecosystem vulnerability based thresholds are present in literature, e.g. an analysis of food webs by Isbell and Loreau (2013) resulted in their discussion statement that complete restructuring of marine food webs can appear when one in four species are affected.

As an alternative for an ecosystem vulnerability based boundary (further referred to as: natural boundary), there is no conceptual argument against expressing the footprint based on boundaries defined in chemical management policies (further referred to as: policy boundary), acknowledging that such a boundary is aimed to be lower than expected natural thresholds. The case studies illustrate the use of both boundary types.

The policy boundary used here is based on the policy criterion known as the 95%-protection level and is derived from a statistical distribution of species sensitivities (Species Sensitivity Distribution, SSD) based No Effect Concentrations (NOEC) derived in chronic ecotoxicity studies, also named the HC5 (Hazardous Concentration for 5% of the species). Because the technical footprint derivation was based on SSDs as derived from acute EC50-data instead of NOEC data, we recalculated the HC5 to an EC50 based value. The technical details can be found in the SI. This resulted in a boundary of 0.1%, which we applied as EC50-based policy boundary equal to the NOEC-based 95%-protection principle.

The natural boundary we applied was derived from aquatic food web models. Mulder et al. (2012) describe how food web structures of fish assemblages in surface waters are affected by different sequences of deletions of species. They described at which fraction of primary deletions (direct local extinction of species caused by the pressure) secondary deletions (local extinction of species caused by the extinction of other species) occurred. We took the level of pressure at which no secondary (indirect) deletions occur in any of the studied food webs as an example of a natural boundary. Given a high variance the lowest value was 3% (Table 5 in Mulder et al. (2012)). In short, the footprint method requires setting of boundaries, for which we used respectively 0.1% of the species affected at EC50 level and 3% of the species deleted as illustrations of a policy- and a natural boundary.

6. Express results as Footprint. In order to derive the footprint from the impact predictions across compartments, impacts in various environmental media at various spatial scales need be aggregated into one number that expresses the overall ecological impact of the production and use of chemicals for the scope. We propose to aggregate the calculated mixture toxic pressures over the compartments for which the footprint is to be calculated (scope) by means of weighting. Two possible ways to do the weighting come to mind, viz. by volume of the compartments or by species richness. Because impact is expressed as potential affected fraction of species in a species assemblage, weighting by species richness would intuitively be a logical first choice. Literature estimates for species richness in marine and fresh water can be found on world level(Dudgeon et al., 2006), but not on the different scales needed for this study, so this option was not applied in the case studies. As second option, good estimates of compartment volumes are available, and these were thus used in the case studies, acknowledging that the higher species variability of fresh water compared to marine water is not taken into account (Goedkoop et al., 2009). Since distribution of substances reaching the ocean will in the first instance be limited to the upper layer, only the first 200-meter (photic zone) of the oceans was taken into account. The volume-weighted mixture toxic pressure (msPAF, % of species) was calculated by:

$$\mathrm{ms}PAF_{\Sigma_s} = \frac{\sum msPAF_s\text{-}Volume_s}{\sum \mathrm{Volume}_s} \tag{1}$$

in which Volumes are the volumes of the compartments on different geographical scales (s).

Given compartment-specific exposures, impacts, a defined boundary, and known compartment volumes, all necessary data are available to calculate the environmental space filled up to the boundary by emissions from a source area, or with other words: the chemical footprint (ChF):

$$ChF = ES_{used} = \frac{msPAF_{\Sigma s}}{msPAF_{max}} \cdot ES_{total} \tag{2}$$

in which $msPAF_{max}$ (% of species) is the maximum acceptable impact (the boundary), ES_{total} is the total environmental space (*here:* km^3) in which the impact takes place (specified under the scope) and ES_{used} the calculated environmental space needed (*here:* km^3) to dilute the calculated impact to a level below the $msPAF_{max}$. For a scope considering the footprint of a single compartment, the formula boils down to the ratio of calculated- and maximum affected fraction, multiplied by volume.

Note that the footprint can easily be converted in a chemical pollution index, by dividing the occupied volume by the available volume, so that index values >1 indicate boundary exceedance.

5.2.2 Case studies.

The method was applied in two case studies, to illustrate technical feasibility, the kind of results it provides and how those results can be interpreted and used.

Organics in Europe. The first case study aimed to derive the footprint of the production and use of mono-constituent organic chemicals in Europe, given the impact of the resulting emissions worldwide, and to derive the relative contributions of these chemicals to the net impact. Emissions of the chemicals during the production, the use and the disposal phase were taken into account, under the assumption that chemicals produced in Europe, are also used and disposed in Europe. Emissions of other substances which may be used in the life cycle of these chemicals (e.g., in their production phase) were not taken into consideration. Production volumes as known from REACH registration dossiers submitted with the European Chemical Agency (ECHA) were used. At the time of this study, dossiers had been submitted for over 6000 different chemical substances. Anonymized emission summary data were used, given data confidentiality issues and the purpose of method illustration. Per April 2011, registration dossiers held data on production and use in Europe of 873 so-called mono-constituent organic substances. Their cumulative market volume was 190 million ton, ranging from 19 million to 1 ton for individual substances. Because reliable toxicity data was not available for all these substances, we focused this study on 630 substances, representing 95% of the total production volume in this category of chemicals. Emissions factors (EF) were tentatively based on the environmental release categories (More detail in SI: http://pubs.acs.org/doi/suppl/10.1021/es500629f).

Pesticides in RMS. The target of the second case was to evaluate the time trend of the worldwide impact of pesticides used in agricultural practice in the combined catchments of the rivers Rhine, Meuse and Scheldt (RMS) by deriving footprints of the use of pesticides in the RMS in 1998, 2004 and 2008, given consideration to all scales. In this case study, emissions during production and transport were neglected and only the use phase of the pesticides was included. Emission data were derived from the Dutch environmental risk indicator for plant protection products (NMI, Kruijne, et al., 2011) and are based on the Dutch purchase figures and per crop usage data regarding plant protection products for 1998, 2004 and 2008. This data includes 287 substances, of which 13 substances could not be taken into account due to lack of toxicity data (less than 2% of the modeled yearly emission volumes). The emissions for the Netherlands were extrapolated to the RMS catchment using the formulae provided by Henning-de Jong et al. (Henning-De Jong et al., 2008). SimpleBox (3.0) was used to estimate the steady state concentrations in the various environmental compartments. The estimated concentrations were compared to observed concentrations at drinking water intake points (REWAB database(Versteegh and Dik, 2012)).

5.2.3 Data

The data sources used for the application of the footprint derivation procedure in the two case studies are summarized in Table 5.2.

Table 5.2 Sources of data used for the case studies described in this paper.

Process step	Data Required	Data sources for case study 1: organic substances	Data sources for case study 2: pesticides
2	Production and use volumes	The European Chemical Agency (ECHA) REACh dossiers	Purchase figures per substance and use patterns per crop(Kruijne R. et al., 2011)
2	Emissions	Emission Factors	Dutch environmental risk indicator for plant protection products (NMI)(Kruijne R. et al., 2011); Extrapolation factors derived from Henning-de Jong(Henning-De Jong et al., 2008) and the CAPRI database(Hierderer, 2012)
3	Physical-Chemical parameters	EPI Suite(USEPA, 2000)	NMI and EPI Suite(USEPA, 2000)
3	Validation of estimated concentrations	Waterbase(Rijkswaterstaat, 2013)	REWAB database(Versteegh and Dik, 2012)
4	Toxicity data	E-toxBase(RIVM) supplemented with data from De Zwart, the ECHA, the European Chemical Substances Information System (ESIS) and ASsessment Tools for the Evaluation of Risk (ASTER) of the U.S. Environmental Protection Agency (EPA).	E-toxBase (RIVM) supplemented with data from De Zwart (De Zwart, 2005), and the university of Hertfordshire (The University of Hertfordshire)

Process step	Data Required	Data sources for case study 1: organic substances	Data sources for case study 2: pesticides
5	Boundary for safe operation	Policy boundary (0.1% of the species affected at EC50 level) and Natural boundary (3% of the species deleted)	Natural boundary (3% of the species deleted)
6	Available environmental space	Volumes used in Simplebox 3.0(Den Hollander et al., 2004)	Volumes used in Simplebox 3.0(Den Hollander et al., 2004)

The degradation constants for water (kdeg,water) are not directly obtained from EPIsuite, but are based on EPIsuite's BIOWIN estimation program and calculated following Rorije et al. (2011):

$$K_{deg.water} = \frac{\ln(2)}{24 \cdot 3600 \text{-} 7300 \cdot e^{-2BIOWIN3}} \tag{3}$$

where $k_{deg,water}$ is the degradation constant for water (s-[1]) and BIOWIN3 is the biodegradability estimate derived with EPI Suite's BIOWIN estimation program. Half-life in soil and sediment are estimated to be respectively twice and nine times the half-life in water, conform Rorije et al. (ibid). All physico-chemical parameters used are included in the SI: http://pubs.acs.org/doi/suppl/10.1021/es500629f.

5.2.4 Uncertainty analysis

Calculated mixture toxic pressures, and hence ChF's, are uncertain as a result of uncertainties in the estimated concentration, median toxicity values (μ_x, EC50) and TMoA-specific spreads in species sensitivities (σ_TMoA, EC50). The uncertainties in the footprints and relative importance's of individual chemicals are derived in two consecutive Monte Carlo simulations. In a first step, the uncertainty of the fate modeling was simulated. This resulted in distributions of plausible concentrations, which were subsequently combined with uncertainties in toxicity data and used as input into a second simulation of uncertain ecological impacts. This sequential simulation of uncertainty in mixture toxic pressure (msPAF) has been described elsewhere (Harbers et al., 2006; Henning-De Jong et al., 2008; Van Zelm et al., 2007). Uncertainty in the toxicity data was described by means of joint normal-gamma distributions of plausible, but uncertain, true values of the toxicity parameters μ_x and σ_TMoA of the various chemicals. Toxicity parameters were estimated by combining the, often scarce, observational information with prior existing information on many other, data-rich chemicals, by Bayesian inference. In this work, we have used an approximation of the so-called empirical Bayes estimation procedure; the formulations used are given in the SI: http://pubs.acs.org/doi/suppl/10.1021/es500629f. Our assessments yielded considerably reduced uncertainty for many data-poor chemicals, which led to less outlier prone estimations of toxic pressures then reported earlier (Harbers et al., 2006; Henning-De Jong et al., 2008; Van Zelm et al., 2007).

5.3 Results

5.3.1 Case 1. Use of organic chemicals in Europe

The resulting yearly emissions to water range from 0.4 million to 0.04 ton per substance on regional RMS-scale and 1.6 million to 0.17 ton on continental scale. Median predicted concentrations and ranges in the fresh and marine waters and oceans are shown in Fig. 5.1A. Comparison of predicted and measured data at the scale of the regional fresh water compartment showed that clear co-variance is poor or non-existent, but that absolute values of predicted concentrations have the right orders of magnitude (Fig. 5.1B).

The calculated mixture toxic pressures on the different scales and their uncertainty ranges are presented in Fig. 5.1C. The highest predicted impacts are expected in regional and continental fresh waters (median values respectively 11.1 and 1.2% species affected at EC50-level). Further away from the emission sources the predicted impact becomes less, due to the combined influence of dilution and degeneration, i.e. 0.2% in regional marine water and close to zero species potentially affected in the continental sea and the oceans. Compared to toxic pressures found in literature, the median toxic pressure in regional fresh water is in line with toxic pressures based on measured values elsewhere (e.g. 8% in river stretches in Ohio (De Zwart et al., 2006)), while the median mixture toxic pressure in regional marine water is lower than reported by Harbers et al. (2006). The latter difference is caused by the inclusions of prior data on ecotoxicity in the uncertainty analysis, resulting in less extreme results in the Monte Carlo simulations.

Acknowledging that impacts exceeds boundaries close to the emission source, we illustrate the integration of the impact scores by volume weighting, and we calculated the European-level ChF by using the policy boundary (0.1% of the species affected at EC50-level). This resulted in a chemical footprint of 7.8x10^3 km^3. This equates to 1.1% of the total EU water volume (Fig. 5.1D). In other words: the virtual water volume filled up to the maximum acceptable limit because of the emissions of organic substances in Europe is 1.1% of the available water volume in Europe (but note the variability of exposure levels at the different scales). When calculated against the natural boundary (3% of the species affected), a footprint of 2.6x10^{11} m^3 (0.04% % of the EU water volume) was obtained.

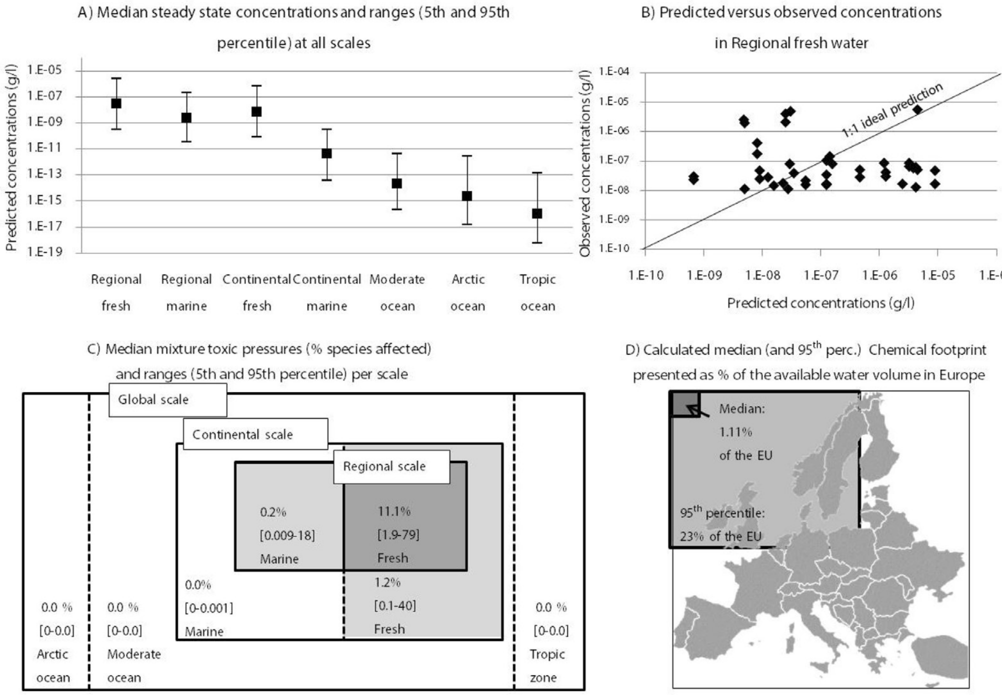


Fig. 5.1. Organic substances case study; A) Predicted median steady state concentrations (and 5th-95th percentile) per scale; B) predicted steady state concentrations versus concentrations observed in the Rhine and Meuse; C) calculated median mixture toxic pressures (msPAF$_{EC50}$, and 5th-95th percentiles), expressed as percentage of species affected at EC50 level. In the dark-shaded area the impacts exceed both the natural (3%) and the policy (0.1%) boundaries, and in the light-shaded areas the impacts exceed the policy boundary only; and D) pollution index expressed in % of production and use of 629 organic substances emitted in the

EU relative to the EU water volume, using the policy boundary.

A contribution to variance analysis showed that, on average, variance in compound toxic pressures derives mainly from uncertainty in ecotoxicity parameters (μ_x and σ_TMoA, both 30%) and to lesser extent from the modeled concentrations (19%), which, in turn, derives mainly from uncertainties in the emissions and degradation rate constants.

5.3.2 Case 2. Use of pesticides in the Rhine/Meuse/Scheldt river catchments

Comparisons between predicted and observed concentrations (Fig. 5.2A) show that both are in the same order of magnitude. The calculated mixture toxic pressures ($msPAF_{EC50}$) declines with the years and is lower at larger geographical scales (Fig. 5.2B). The resulting footprints of respectively 1998, 2004 and 2008 after weighting and footprint calculations (EC50 based; natural boundary used, km3) are summarized in Fig. 5.2C and compared with the available environmental space in the RMS catchment (km3, dashed line in the chart). The analyses suggest a declining footprint for the virtual 'RMS average water', over time, stabilizing in more recent years. The decline in the modeled impact is due to a reduction in quantity of pesticides used per ha and a switch to pesticides with other properties, and not for example to a decline in agricultural area. This pattern is in line with other evaluations on the use of pesticides in the Netherlands (De Snoo and Vijver, 2012; Henning-De Jong et al., 2008; Struijs et al., 2010; Van Eerdt et al., 2012) and Europe (epp.eurostat.ec).

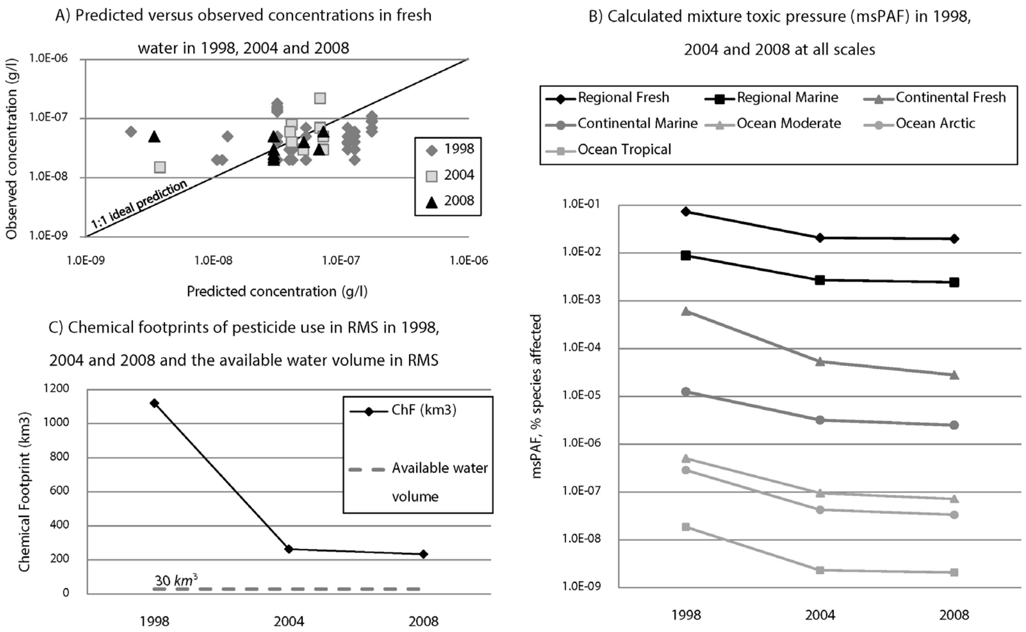


Fig. 5.2 Pesticides case study; A) predicted steady concentrations versus concentrations observed at drinking water intake points in the Rhine and the Meuse; B) Calculated mixture toxic pressure (msPAF, % species affected at EC50 level) at different scales due to pesticide use in the Rhine Meuse Scheldt catchment (RMS) in the years 1998, 2004 and 2008; C) The chemical footprints (natural boundary based, 3% of the species affected) of the pesticide use in RMS in 1998, 2004 and 2008, given the impact on all scales, expressed in km3

and set against the available environmental space (ES): the water volume in RMS.

5.3.3 Relative contributions of individual substances

The individual substances can be ranked according to their relative contribution to a ChF. Existing procedures to rank substances regarding their environmental hazards can be found in literature, and are based on chemical properties(Strempel et al., 2012), chemical properties and use characteristics(Mitchell et al., 2013) and multi-media model approaches (Arnot et al., 2012). The latter approach was adopted in this study as follows:

$$R_x = 1 - \frac{Ch_{n-x}}{ChF} \tag{4}$$

with R_x being the relative contribution of substance x and ChF_{n-x} the Chemical Footprint without the contribution of substance x.

In both case studies, Monte-Carlo simulated margins of uncertainty of R were large. The 5th percentiles of R was lower than 0.1% for every substance. Adopting the R-value of 10^{-3}, suggested by Harbers et al. as a limit of "significant contribution", it can be concluded that it is not certain for any of the substances if their individual contribution to the ChF is significant. Also, the 95th percentile of the R was lower than 0.1% for most of the substances (organic substances case study: 75%; pesticides case study: 85%). Thus, most individual substances are not likely to contribute significantly to the net toxic pressure, considering their spread over Europe's total water volume. It follows from the uncertainty analyses that from the remaining substances (resp. 25% and 15% of the substances) only a small fraction of the chemicals (6% of all the organics and 3% of all the pesticides) account for the larger fraction of the total toxic pressure on aquatic ecosystems. In spite of the meaning of this outcome pattern that only few chemicals contribute to impacts, it cannot be determined with certainty which of the chemicals belong to the contributing group at the studied scale due to the wide confidence bands. This is not so much a weakness of the method – it relates to uncertainties in- and lack of data and the scale of study. The results are comparable with results found by De Zwart (2005). They found that 95% of the predicted risk due to pesticide use in the Netherlands is attributable to only 2.7% of the substances.

5.4 Discussion

A versatile and comprehensive novel method has been developed that provides a quantitative measure describing the environmental space (here: water volume) needed to dilute chemical pollution due to human activities to a level below a specified impact boundary condition. The method proceeds beyond the works of Harbers et al. (2006) and Henning-de Jong et al. (2008) by expressing mixture-impact related outcomes in terms of a chemical footprint, in line with general requisites for footprints (Čuček et al., 2012), and considering an explicitly defined set of optional boundaries from the realms of chemical management practice (policy boundary) and research on ecosystem vulnerability (natural boundary). The method combines years of experience in the field of fate and effect modelling, with contemporary developments concerning footprinting and planetary boundaries, and adds novelties on different levels, e.g. the derivation and application of two types of boundaries for chemical pollution, the way uncertainty analysis is designed and the use of up to date data from policy dossiers. The method is flexible in the sense that it can be adjusted and applied on different scenarios and assessment targets, ranging from assessing the chemical footprint of single-chemical emissions from a point source or product to nearby water bodies up till world-wide assessments from regional use of multiple chemicals, as illustrated in both case studies. Finally, the method consists of a logical, robust sequence of analytical steps that allows for validation of outcomes of the separate steps, it technically results in outputs of the desired kinds, as illustrated in two case studies, and the results of those case studies are in line with other data, while appearing to be meaningful. The results of ChF assessments can be used to evaluate management of chemicals on different spatial levels in the light of set boundaries, to support decision-making by exploring the ChF of alternative scenarios and to support prioritizing of chemicals or sectors or areas to focus on. A need recently voiced in literature (Bu et al., 2013; Giubilato et al., 2014).

The first application of the method in two case studies has highlighted that individual steps can be improved regarding quality of input data and the subsequent modeling steps, such as the emission estimation and uncertainty assessment, and/or further discussed, such as the used boundaries and the weighting method. Given uncertainty bounds in the results of the current case studies, we prefer exploring relative outcomes (policy-evaluation, prioritization, etc.) rather than absolute outcomes. We address various influential aspects specifically below.

First, the uncertainty in the ChF and R calculations stems from uncertainties in the substance-specific toxicity value (μ_x), the TMoA-specific spread in sensitivities of species (σ_TMoA), and in the modeled steady-state concentrations. Uncertainties in the steady-state concentrations in turn originated largely from the uncertainties in emissions and degradation rate constants. Use- and substance-specific emission factors are usually hard to obtain and uncertain (Arnot et al., 2012; Warmbaugh et al., 2013). This was also true for the two case studies in this study. Because the purpose of the case studies was testing and improving the outlines of the method, rough emission estimates sufficed. To improve the accuracy of chemical footprints obtained with our calculation procedure, research on derivation of emission factors is recommended. Uncertainty in the toxicity parameters μ_x and σ_TMoA could be reduced by including more toxicity data per substance, more substances per TMoA and improved TMoA assignments. Van Zelm et al. proposed research on the use of Quantitative Structure-Activity Relationships to supplement experimental toxicity data (Van Zelm et al., 2007) and also the use of prior data in these assessments should (and will) be further investigated.

Second, the application of a multi-media mass balance model (SimpleBox) and the two biological models (concentration addition and response addition) to estimate predicted ecological impacts of mixtures of substances comes with simplifying assumptions. Such assumptions need be made explicit and detailed to support the interpretation of the final results. The box-modeling approach assumes environmental compartments to be spatially homogeneous. In reality impacts will be lower in the larger sub-compartments of the environment and higher at so-called 'hot-spots'. Instead of box models, spatially resolved models could be used, which would give more insight in hot spots, but simultaneously make the analysis more data demanding and the aggregation more difficult. Further, the method can also proceed based on measured concentrations. Likewise, variability in time, which is shown to be relevant for e.g. pesticides by de Zwart et al. (2005), is neglected. Also, the case studies did not treated bioavailability; the model as applied assumed dissolved concentrations in water are fully available to the aquatic species. This can be added to the method, which is of particular importance for e.g. metals, as shown by Gandhi et al. (2010). Finally, the mixture modeling asks for identifying the toxic mode of action or mechanism of action per substance. However, the appropriation of TMoA per substance varies amongst different sources of toxicity data used in this study and was missing for many substances. Applying a mode of action needs detailed information on the substance at stake, which is not always readily available, especially in the case of organic substances. Furthermore, substances can have multiple Mode of Actions, which makes categorizing substances complex. Despite this, quantitative mixture impact predictions may be considered robust (Drescher and Boedeker, 1995).

Third, reconsideration of the scopes of the case studies reveals important aspects. When working on (results of a) ChF analysis one must always be aware of what has actually been analyzed. In our case we focused on the impact on aquatic ecosystems. The method could be applied to include terrestrial and aerial impacts, but because of the (relative) lack of ecotoxicity data for these compartments they were excluded from the study. Next to impacts on other compartments, other types of impacts could in principle be included in the method. For example SSDs for ozone exposure of plant species have recently been developed (Van Goethem et al., 2013), so that an ozone footprint may be derived. In addition, it would be interesting to apply the method on risks of mixtures of substances for adverse impacts on humans.

Fourth, weighting is a key step in the method. The application of weighting to aggregate the impacts on different scales and compartments needs to be explicit, and could be (and is) disputed, because aggregation implies information loss. However, policy usage often implies that information is presented in an easy to interpret and communicate way(OECD, 2002). The latter explains the success of footprint methodologies, despite the suite of methodological questions and suggestions raised in literature (Böhringer and Jochem, 2007; Fiala, 2008). Next to loss of information, aggregation may lead to results with higher uncertainty ranges (Heijungs et al., 2003). That is why the aggregated results should always be accompanied by the underlying intermediate results and explicit descriptions of the assumptions and aggregation methodologies applied.

Of particular interest is, fifth, the issue of boundaries. Our approach presents two operational methods to define this key item for the footprint method for the first time. There is substantial latitude to further discuss and derive boundaries, in line with the general issues discussed in the Planetary Boundaries papers, ongoing specific (policy) discussions on safe single-chemical boundaries, and current works on biodiversity impact boundaries. The complexity of natural systems should be acknowledged (Brook et al., 2013), and would suggest a distribution of spatially variable, and ecosystem specific natural boundaries. Some unifying principles needs be chosen, which could build on analyses of known tipping- or bifurcation points for local ecological processes, such as lake eutrophication or loss of semi-arid vegetation due to dryness (Scheffer et al., 2009) or the human induced restructuring of marine food webs(Isbell and Loreau, 2013). The natural boundary used in this paper was taken from a study that clearly shows variability of secondary species deletions across food webs. The idea of natural boundary diversity based on food web collapse will be further explored, within the context of both food web structure and function. The so-called ecosystem services concept can be a unifying principle here, implying that boundaries are derived in relation to services offered by ecosystems and their processes to man. Applied to all global emission sources and combined with a distribution of natural boundaries conditions or a uniform policy boundary our method would yield a proposal to define the global boundary for chemical pollution as discussed by Rockström et al. (2009).

Finally, it should be noted and clearly communicated that the resulting footprint, or: environmental space (water volume) needed, is hypothetical. Critics state that this could be easily misinterpreted (Van den Bergh and Grazi, 2013), especially because the hypothetical environmental space is compared with the actual environmental space available to express the sustainability of a system, neglecting gradients from emission source to distant compartments. When comparing the footprint (volume needed) with the environmental space available (volume present), the latter should be obtained and communicated with care. A visualization of this point can be found in the SI (Table S4, http://pubs.acs.org/doi/suppl/10.1021/es500629f). From an environmental (biocapacity) perspective, derivation of the available environmental space of a bioregion (e.g. river basin) makes more sense than that of a political region (e.g. country) (Van den Bergh and Grazi, 2013).

The points discussed so far do not invalidate the foundations, logics and versatility of the proposed method. Rather, the issues define where there is latitude for technical improvements of the various steps, whilst our analyses suggest that the final results of the presented footprint method already in their current status can very well be used, i.e. to support priority setting in chemical management (both case studies), or in the evaluation of the success of measures taken by risk managers (second case study).

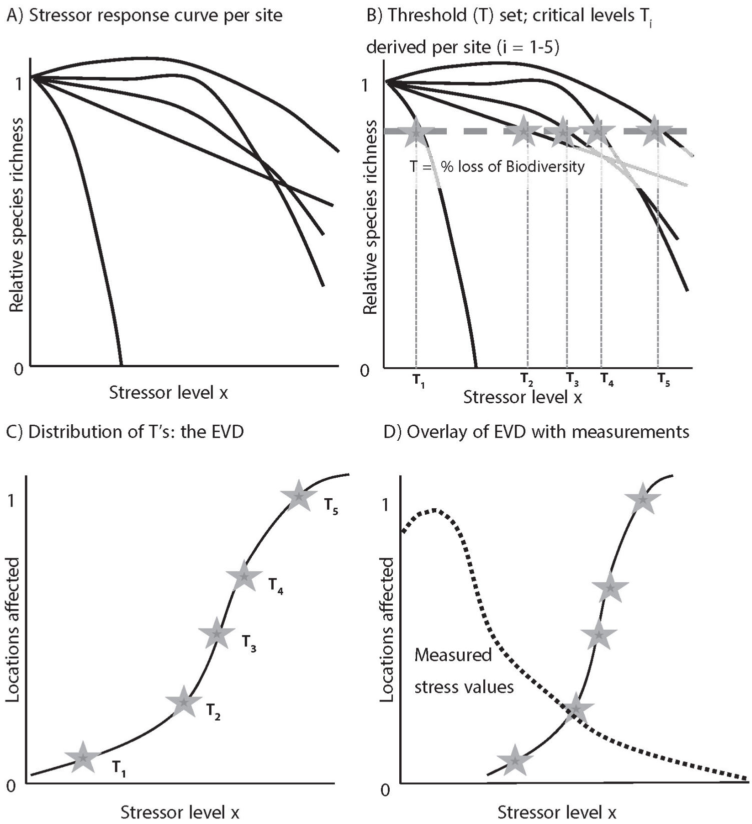
5.5 Acknowledgements

Funding by the Strategic Research Program of RIVM is gratefully acknowledged (project number S/607020). We thank M. Huijbregts, J. Struijs, A. Hollander, D. De Zwart, R. Heijungs and C. Mulder for fruitful discussions and A.M.A. Van der Linden and H. Van Wijnen for support with data preparation.

6 Identification and ranking of environmental threats using ecosystem vulnerability distributions

Authors: Michiel C. Zijp, Mark A.J. Huijbregts, Aafke M. Schipper, Christian Mulder, Leo Posthuma

Submitted for publication



Abstract
Responses of ecosystems to human-induced stress vary in space and time, because both stress and vulnerabilities to that stress vary in space and time. Presently, impact assessments mainly take into account variation in stress levels. Scientists involved in environmental impact assessment and management struggle with designing vulnerability assessments that are sufficiently operational. Moreover, assessment of some stressors is disjunct from others, with toxic chemical mixtures as an example. Our method for deriving ecosystem vulnerability distributions (EVDs) provides a new way to gain insight in human-induced impacts in a region. The method acknowledges spatial variation of both multiple stressors and ecosystem vulnerabilities and bridges the gap between toxic chemical mixtures and other stressors. EVDs can be derived based on a selection of locations, existing monitoring data and an environmental boundary. A case study on Ohio's freshwater ecosystems showed that the approach is feasible and allows stressor identification and ranking. Physical habitat characteristics and nutrient loads ranked highest as current stressors, i.e., implying species losses higher than 5%. Deriving and using EVDs is an operational way to complement the common approaches to account for variation in the stressors with a novel approach to account for variation in the stressed.

6.1 Introduction

Ecosystems are under influence of human induced stress all over the world (Maxwell et al., 2016). Some authors even speak of a nearly global 'anthropogenic biosphere' (Ellis, 2015). The diagnosis of environmental impacts, including identification and ranking of stressors, is of key importance for protective and restorative management. The impact of human stress on ecosystems depends not only on the types and magnitudes of the stress, but also on the vulnerability of the receiving ecosystems to that stress. Hence, to reach policy goals regarding sustainable management of our environment, such as the UN development goals (UN, 2015a), the European Water Framework Directive (EC, 2000) and the (American) Clean Water Act, knowledge is required on both the types and magnitudes of stress we pose on ecosystems as well as on ecosystem vulnerabilities. Ecosystem vulnerability shows large spatial and temporal variability (Brook et al., 2013; Simkin et al., 2016; Steffen et al., 2015; Strona and Lafferty, 2016), i.e. it depends on the interplay between environmental conditions and species assemblages, which are both location-specific features (De Lange et al., 2010; IPCC, 2001; Ippolito et al., 2010; Mulder et al., 2012; Posthuma et al., 2014; Thywissen, 2006; Villa and McLeod, 2002). Quantification of variation in ecosystem vulnerability within a region is complex for various regions. Many definitions exist on what vulnerability should encompass, most definitions are not operationalized (Beroya-Eitner, 2016), and when they are operationalized they can be very data intensive. A complicating factor is that in most situations ecosystems are simultaneously exposed to different stressors (Goussen et al., 2016). Addressing these different stressors in a single assessment is a challenge, given the current mono-disciplinary method development and communication, e.g. in the case of chemical pollution(Bernhardt et al., 2017). As a consequence, variation in stress is commonly handled in environmental management, but variation in vulnerability is not.

Neglect of the latter however, hampers assessments and management of human-induced impacts on the environment, restoration activities, and decisions regarding landscape planning (Simkin et al., 2016; Steffen et al., 2015; Vaquer-Sunyer and Duarte, 2008). As an example, regulations on the production and use of chemicals (EC, 2006) are based on the response data collected for of a small selection of individual species under single-stressed laboratory conditions and not on knowledge on the variability in ecosystem vulnerability of actual occurring species assemblages under multi-stressed conditions in the field (Vindimian, 2001). As a consequence this neglect of ecosystem vulnerability variation leads to cases of environmental under-protection (Schäfer et al., 2007; Vaquer-Sunyer and Duarte, 2008) and over-protection (Beketov et al., 2013).

Related to the concerns on environmental impacts of man-made pressures many countries invest in monitoring programs that measure biological and environmental quality. This allows for a responsive management approach when the monitoring results show an aberrant (place) or decline (time) in biological quality. Here, we show how such monitoring data can be used to not only assess variability of stressors, but also of ecosystem vulnerability within a region. And, furthermore, how the collated knowledge on stress and vulnerability variations can be used to improve the stressor identification and ranking for a region.

In this study ecosystem vulnerability is operationally defined in relation to the level of stress imposed by a stressor (an abiotic characteristic) to reach a natural or selected adverse impact level; the higher the stressor level needed to induce such an impact level, the less vulnerable the ecosystem for that stressor. The ecosystem vulnerability is assumed to depend on the vulnerability of the site-specific assemblage of species and environmental conditions (Beroya-Eitner, 2016; De Lange et al., 2010; IPCC, 2001; Ippolito et al., 2010; Thywissen, 2006; Villa and McLeod, 2002). Therefore, ecosystem vulnerability will vary in a region with varying combinations of species and environmental conditions. This definition is made operational for environmental assessment and management with a method called the ecosystem vulnerability distribution (EVD). The idea of the EVD is to quantify the range of ecosystem vulnerabilities of all sites in a region based on the assessment of a selection of sites that are relevant for the assessment, e.g. reference sites. This is similar to the principles and uses of species sensitivity distribution (SSD) models, which describe the range of sensitivities of all tested species for a chemical based on experiments with a selection of species (Posthuma et al., 2002). The analysis of the overlay of the distribution of species sensitivities with a distribution of exposure to a chemical has had, since the groundbreaking studies on atrazine risks across North America (Solomon et al., 1996), major influence on environmental protection, assessment and management of plant protection products on the global scale (Solomon et al., 2013). As will be shown in this manuscript, EVDs can likewise be applied, but now for stressor identification and ranking across multiple potentially relevant stressors using field data on environmental conditions and location specific assemblages of species. As such, the approach also bridges the gap between assessments of the impacts of chemical mixtures and those of other stressors (Bernhardt et al., 2017).

Currently, impacts of environmental stress on species and assemblages can be delineated with various modelling techniques, each focusing on a component of vulnerability. For example, at the species level, the species sensitivity distribution approach provides insight in the inter-species variability in sensitivity for a stressor (Posthuma et al., 2002), while species distribution models (SDMs) relate the occurrence or abundance of individual species to environmental conditions (Guisan and Thuiller, 2005). At the assemblage level, food web modelling provides insight in the interaction between species, which helps to understand and predict direct and indirect impact of stressors for individual food webs (De Laender et al., 2015; Mulder et al., 2012). At the assemblage level, moreover, stacked species distribution modelling (sSDM) allows combining the SDMs of local species to an assemblage-level metric in order to describe and predict response to stress at the level of species assemblages(Schipper et al., 2014). The latter approach, combined with the concept of distributions as applied to model species sensitivities, formed the basis to derive ecosystem vulnerability distributions (EVD). These modelling approaches combined in an EVD allow for evaluating context-dependent responses of species assemblages to stress. It takes into account heterogeneity in environmental conditions in the region under consideration and it covers both various stressors and multiple species.

The EVD for a region and a selected stressor is derived in four steps, depicted in Fig. 6.1. The derivation of an EVD starts from biomonitoring data, the derivation of stress-response models for species and assemblages from those data (SDM and sSDM), and the use of a selection of minimally-disturbed locations in the region as reference sites (Stoddard et al., 2006) (see Methods). To show the methodological principles (and versatility) of EVDs and their utility for stressor identification and ranking, the method was applied on a case study with assemblages of fresh water fish species across Ohio (U.S.A.). Fresh water ecosystems are one of the most human impacted habitats (Xiong et al., 2016), and Ohio is a state with a wide variety of human activities and environmental conditions, while the State has collated one of the richest monitoring data sets currently available with close links to the States' environmental assessment and management policies.

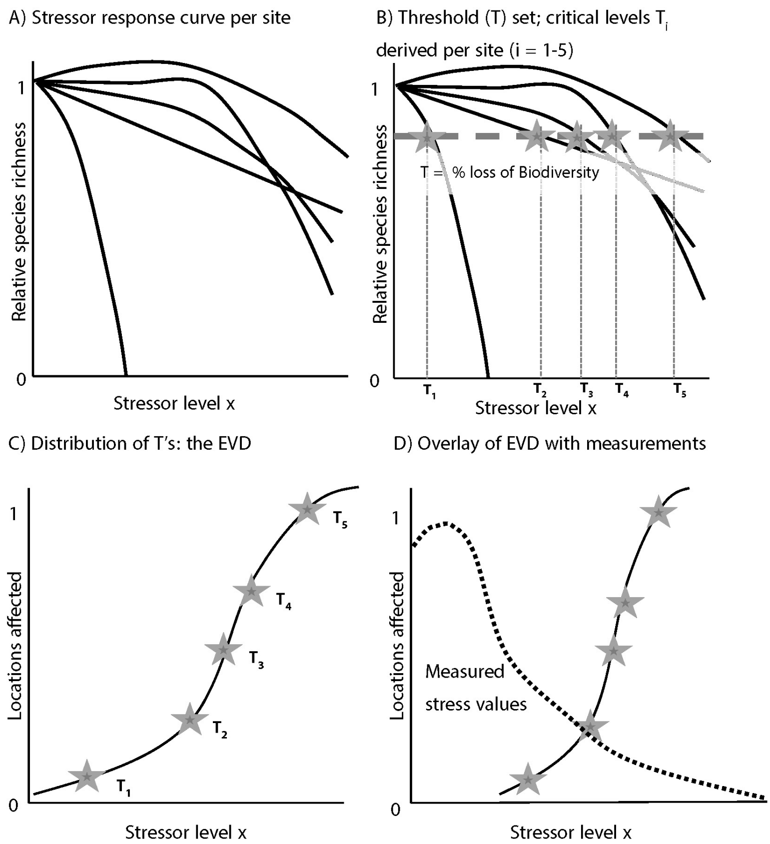


Fig. 6.1 Derivation of ecosystem vulnerability distributions (EVDs) using five sites relevant to the assessment problem as reference sites. To obtain the start position shown on panel A, region-wide biomonitoring data are used to derive species-specific stress-response models (SDMs), which are subsequently collated on the basis of knowledge of species occurrence at the five reference sites in the region of interest (sSDM), and thereafter normalized to a value of 1 for the reference condition. The relative species richness (Y) as a function of a stress level is the species richness at stress x divided by the species richness at the reference situation. To derive the EVD, for a stressor the approach is as follows: A) the change of the relative species richness in relation to a stressor for each reference site is modelled in silico using the stressor-response models (other stressors kept constant); B) a threshold (T, expressed as a reduction of Y) is chosen to enable operational characterization of vulnerability differences across the reference systems for a stressor, here: representing a certain loss of biodiversity; and for every reference site (i=1-5 in this example) the stressor level that results in impact=T is quantified (yielding threshold stress levels Ti); C) the distribution of thresholds Ti=1-5 across the reference sites yields the distribution of vulnerabilities (the EVD) for that stressor in the region of interest; D) the EVD is used to identify if the stressor is a problem in the region by comparing the distribution of the measured stressor values of the sampling sites in the region with the EVD. Overlap between the two distributions signals that stress exposure transgresses the vulnerability, implying stressor effects. The presence and degree of overlap for different stressors can be interpreted in terms of stressor identification and ranking.

The example case study started with a collection of biomonitoring data for Ohio freshwater systems, deriving species-specific SDMs for the occurrence of fish species in the studied region and selecting a subset of 18 sites that are in good and stable ecological condition and represent the region under consideration (further: reference sites). The environmental stressors taken into account in the fitting of the SDMs were water acidity (pH), mixture toxic pressure, nutrient loads, hardness, conductivity and habitat quality. The latter was quantified by the Qualitative Habitat Evaluation Index (QHEI (Ohio EPA, 2006)): a measure for the physical-habitat quality for fish, developed for running waters in Ohio that ranges from 0 (poorest quality) to 100 (maximum quality). Mixture toxic pressure was expressed as the multi-substance Potentially Affected Fraction of species ($msPAF_{EC50}$) derived from the 50% effect concentrations of metals, ammonia and nitrite (De Zwart and Posthuma, 2005; Pilière et al., 2014). The metric 'stacked probability of occurrence' (SPO) is a powerful indicator for species richness (Schipper et al., 2014), and is used as metric for relative species richness (as in Fig. 6.1). Therefore, estimations for species richness (SPOs) were derived for the local stressor levels and local fish species assemblage for all reference sites. With these SPOs, and after their normalization, stressor-specific response relationships of relative fish species richness are derived for each reference site by changing in silico the value for one stressor while maintaining the observed reference field conditions for the other stressors (Fig. 6.1A). In this step the hypotheses that ecosystem vulnerability within a region is a distribution rather than a single fixed value is already substantiated when the modeling yields different stressor-response curves across sites. Next, derivation of the ecosystem vulnerability distributions requires a (chosen or natural) threshold (T) for operationally judging a stressors' impact, e.g. 5% fish species loss (Fig. 6.1B). Then, the stress-levels at which this impact is reached at the different reference sites are derived and summarized as a distribution (Fig. 6.1C): the ecosystem vulnerability distribution (EVD) of that stressor. The stressor identification and ranking steps are finally made by evaluating the overlays between the distribution of the stressor levels of a region (all data) and the associated stressor-specific EVDs. In short, the EVD translates the hypothesized across-site ecosystem differences in vulnerability into an operational metric to identify and rank stressors that threaten the biodiversity in a region.

6.2 Results

6.2.1 Species-specific responses to stress

The species distribution models (SDMs) of all but five species had significant coefficients *(c, or both c and c',* see Methods, §6.4) for all stressors, which means that the occurrence of most fish species co-varies with all stressors. The stressor-specific coefficients, however, varied widely across the species, which means that different fish species respond differently to the same stress. This reflects our hypotheses of the importance of EVDs: the species-assemblage response to stress is a location-specific feature. The SDM coefficients per fish-stressor combination can be found in the Annex, §6.6 (Table S6.3 and summarized in Figure S6.1). Looking to the SDM-coefficients in more detail revealed that drainage area and QHEI have in general (median values, Fig. S6.1, grey bars) the most positive association to the probability of occurrence of fish species compared to the other stressors. TP has the smallest median coefficient, and thus the most negative influence on the probability of occurrence of most fish species. Closer inspection also revealed that the 5th percentile of the estimated coefficients across the species is negative and the 95th percentile is positive for all the stressors (Fig. S6.1 left and right side bars). This means that for all stressor an increase in the value of a stressor implies positive as well as negative changes in the probability of occurrence of different fish species, depending on the species. Finally, the SDMs described the species occurrence well: 90% of the SDMs have an AUC > 0.7 (see Methods on AUC). Thus, the results of the SDM-analyses support the expectation that ecosystem vulnerability is distributed and that an EVD can be derived, as the SDMs are of sufficient quality (technically relevant) and show diversity (ecologically relevant).

6.2.2 Assemblage-specific responses to stress

The normalized stacked probability of occurrence (SPO) was derived as impact metric in the stressor-response analyses. As expected, predicted SPOs correlated well with observed species richness data (Fig. S6.2 at §6.6). The combined qualities of the SDMs and the SPO are interpreted as a good basis for the subsequent step of stressor-response modeling. The evaluations of the changes in SPO per site with an in silico change in one of the stressors resulted in three key observations. First, assemblages of fish species at different reference sites respond differently to imposed stress (Fig. 6.2), whereby the 'points of departure' (the conditions at the reference sites, plotted as dots in Fig. 6.2) illustrate the variability of the natural conditions for the set of reference sites. Second, different stressors yielded stressor-response curves of different shapes. Changes in pH resulted in bell-shaped curves, indicating that both an increase and a decrease of the stress level can cause SPO decline, with – for part of the sites – an increased predicted SPO with slightly lowered pH. Mixture toxic pressure and hardness showed a skewed optimum: a slightly increasing mixture toxic pressure or hardness yielded (slightly) increased SPO, with more significant SPO reduction at increased stressor values. The response curves for QHEI and TP are sensitive and approximately linear for most sites. Change in conductivity results in only a marginal decline in SPO at 15 of the 18 sites. This observation resulted from the counterbalancing effect of the fact that various species are predicted to respond positively and others negatively to an increase in conductivity. Change in TN showed a diversity of responses: SPO at some locations respond linear, others show an optimum-, or a skewed optimum. Third, and key for the next steps, the type of analysis does not yield values for natural threshold of impacts, as all species-loss curves are gradual (Brook et al., 2013).

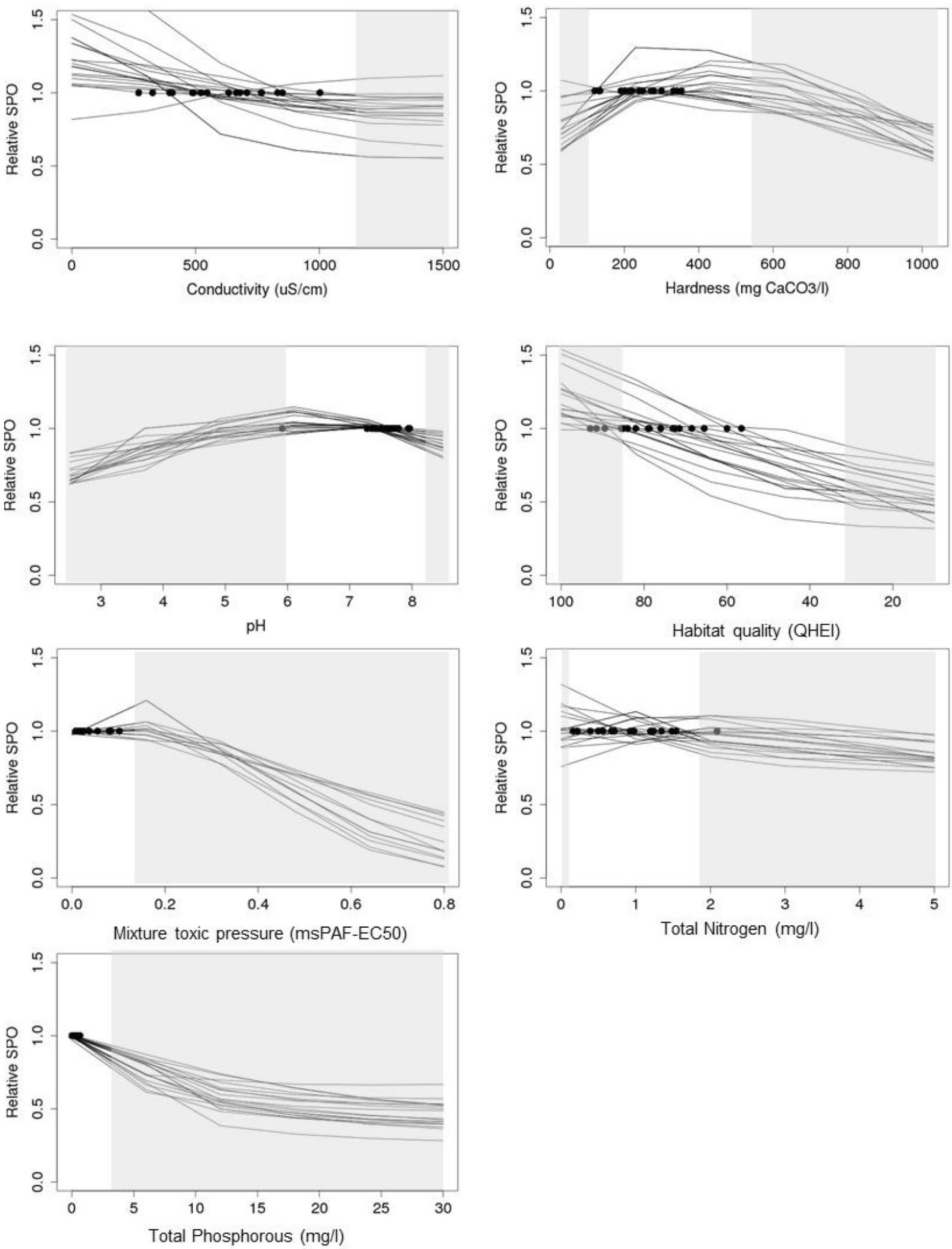


Fig. 6.2 Modelled stressor-response relationships for Ohio fish species assemblages for the selected 18 reference sites for conductivity, hardness, pH, QHEI, mixture toxic pressure, Total N and Total P (x-axes) and the relative SPO (Y). Y-values higher than 1.0 indicate an increase in SPO compared to the observed situation at the reference site and vice versa. The 'points of departure' – SPO at reference conditions – are shown as black dots. The white areas contain 90% of the stressor-level variability across Ohio.

6.2.3 Ecosystem vulnerability distributions (EVDs)

Predicted assemblage-level responses to stress were indeed specific for sites and stressors (Fig. 6.3) Due to the low response for conductivity and the complex response for TN, example ecosystem vulnerability distributions were not derived for these two stressors. For pH two EVDs were derived: one for increasing stressor values and one for decreasing values, as compared to the reference-measured value. The point-values that define the ecosystem vulnerability distributions in the Figure were derived with a selected impact threshold T=5% SPO-reduction, but the observed patterns are robust for other selections of T (Annex Fig. S6.4). This suggests that the value-choice for T – needed in the absence of a natural threshold effect (Brook et al., 2013) – does not affect the conclusion that the vulnerability for a stressor varies across ecosystems. The most vulnerable and least vulnerable reference sites, and the pattern, determine the EVD range and its meaning for stressor identification and ranking. For example, the stress levels that would induce 5% SPO reduction due to mixture toxic pressure range between 0.11 and 0.25 across the sites. That is, the mixture toxic pressure for the least vulnerable system must be more than twice as potent as for the most vulnerable site to cause 5% SPO-reduction.

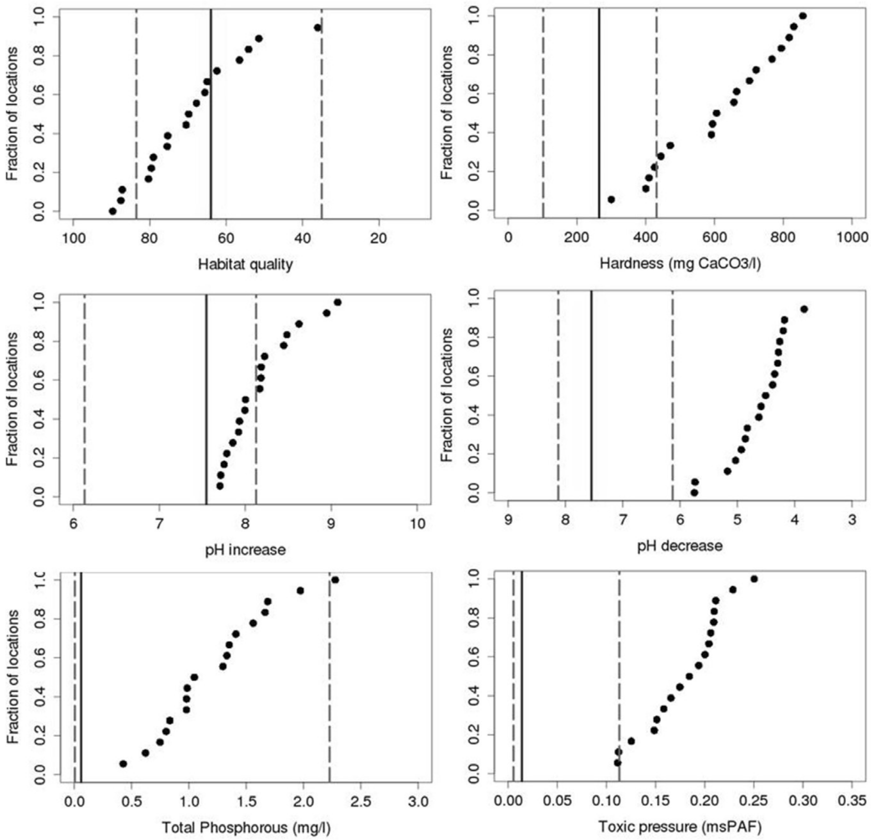


Fig. 6.3 The ecosystem vulnerability distributions for hardness, habitat quality, pH (decrease and increase), mixture toxic pressure and total phosphorous, with on each x-axis the stressor levels and on the related y-axis the fraction of reference sites at which the threshold (5% species loss) is transgressed. The median stressor values across Ohio are indicated with a vertical solid line. The areas between the two grey lines contain 90% of the stressor-level variability across Ohio.

6.2.4 Utility: landscape-level stress identification and ranking.

The stressor identification and ranking now follows from the overlay over the measured stressor data and the EVD, considering the region from which the reference sites were selected and which they aim to represent. The degree of overlay of the EVD-ranges with the ranges of observed stressor values across the 1,826 monitoring sites of Ohio can be interpreted as a landscape-level stressor identification and ranking (Fig. 6.4), similar to the approach in ecotoxicology where the overlay of exposure- and sensitivity distributions for species yields a powerful summary of ecotoxicological risks of a chemical for an area, with proven use for practical decision support purposes in the management of hazardous chemicals (Aldenberg et al., 2002; Solomon et al., 1996; Solomon et al., 2013). Note that this study shows that the problem of often distinct assessments for chemical pollution and the other stressors is bridged by combining them in the derivation of EVDs (Bernhardt et al., 2017).

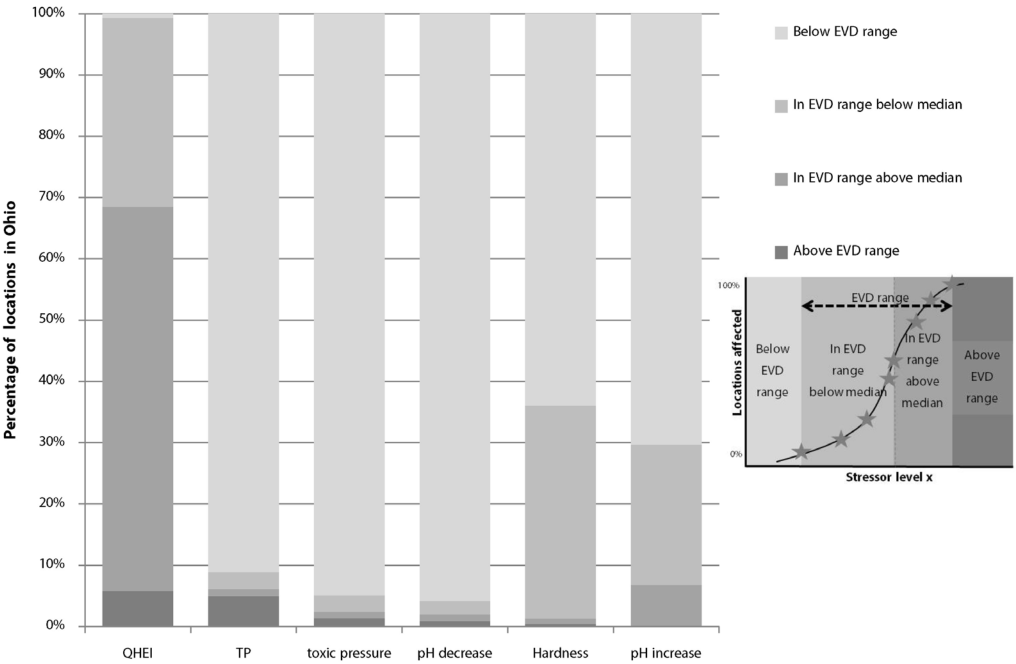


Fig. 6.4 The percentage of locations exposed to a stressor at levels below the EVD range are the lightest shade of grey and the percentage of locations exposed to stress above the EVD range are the darkest shade of grey. The two shades of grey in between are the percentages of locations exposed in the EVD range: light grey below the EVD median and dark grey above EVD median.

Ecological impacts in the study region (defined by 5% reduction of SPO) are mainly related to QHEI and TP. The comparison of the EVDs with field observations (Fig. 6.4) shows that of all stressors, reduction of QHEI is of the highest concern for the occurrence of fish species. At nearly 6% of the locations the current QHEI results in a SPO reduction higher than the threshold T of 5% (dark grey bar in Fig. 4). Furthermore, it holds for 94% of the locations in Ohio that they are at risk for transgressing T-5% for QHEI. For the other stressors, most of Ohio is exposed to stressor values below the EVD range and these sites are thus not at risk at the level of 5% SPO reduction. Next, for TP 91% of the region is not at risk of 5% SPO reduction. Of the other 9% of the Ohio sites, 5% is exposed to TP concentrations above the EVD range implying a higher loss of fish species than 5% due to increased levels of TP, and 4% of Ohio is exposed in the range of the EVD stress levels and are thus at risk of more than 5% loss of fish species due to increased levels of TP. These stressor identification and ranking results related well to other methods of stressor identification and ranking, both reported as well as confronted with long-term experience with Ohio fish assemblages and man-made stress (pers. comm. S.D. Dyer): habitat modification and nutrients were identified amongst the five top stressors for freshwater ecosystems in the recent Integrated Water Quality Monitoring and Assessment Report of Ohio Environmental Protection Agency (Ohio EPA, 2014). This is not only a cross-confirmation of current outcomes. The EVD-method – in addition – allows for a dynamic assessment of implications of future changes in stressor levels caused by societal developments and/or stressor abatement strategies, expanding on the static situation analyses that are more common. This links ideally with the recent call for solution-focused approaches, with are characterized by attention for alternative solutions and the need to explore the expected responses of alternative management scenarios (Schäfer et al., 2016; U.S. NAS, 2009; Zijp et al., 2016).

6.3 Discussion

6.3.1 Methodological robustness

The shape and position of EVDs depend on methodological choices, which may imply choice-dependency of interpretations, an undesired potential method feature. We investigated robustness of our conclusions for various alternative choices. Regarding the fit of the SDMs, the outcomes where robust for choices in SDM fit preferences. Regarding the choice of reference sites, the example case study shows particular consequences of choice: the reference sites were based on allometric relationships amongst the species in the local food webs, which resulted in a selection of stable systems with sites that includes locations with human influence, i.e. activities upstream. This is why e.g. QHEI is only 56 and 59 at two of the locations. Apparently, human activity and stable fish assemblages can coexist. This also means that the probability of the assemblage to improve integrity can be increased by reducing the level of such a stressor, which is indicated by a relative SPO higher than 1.0 in Fig. 6.3. The EVD-method can, however, also be applied on a selection of impacted sites, for example to analyze the vulnerability distribution (its position and shape, as well as the overlay with the landscape stressor distribution) of a set of sites more influenced by human impacts. That would provide insights whether stressor levels would further affect the species assemblages at such sites. Regarding the chosen response metric, the choice for a higher T results in a distribution that is shifted to the right, with (in this example) a larger difference in vulnerability between locations (Annex Fig. S6.4). Thresholds might also represent a natural 'breakpoint' in stressor-response curves, also called tipping points (Scheffer et al., 2009). A possible natural T would be the cascading stress level at which an indirect (secondary) species deletion is predicted to occur (Mulder et al., 2012). All these findings do not invalidate the main outcomes of this study, that vulnerability varies across sites and that the overlay of an EVD with stressor data supports stressor identification and ranking.

Any ecosystem depends both on the vulnerability of individual species within its boundaries as well as upon the interactions among all the occurring species and the present condition of the environment. The method presented in this paper does not yet include inter-species interactions. Examples of SDMs that include inter-species interactions for a large assemblage of species are not available, as they are statistically impossible to describe. This is different for few-species systems, e.g. where species have strong mutual interaction like between a typical plant species and an associated pollinating species (Pellissier et al., 2013). Due to competition for food resources, adding trophic data to the occurrence of fish species might improve the SDMs in the near future when needed (Wisz et al., 2013), although calculation time increases exponential with every extra stressor.

6.3.2 Implications

We showed that ecosystem vulnerabilities for a stressor in a region are distributed, in addition to the well-known distributions of stressor levels in a landscape. Our findings primarily support stressor identification and ranking. But they also imply, amongst others, that transgression of a protective criterion, such as a concentration criterion for a chemical (e.g. the Clean Water Acts' 'Ambient Water Quality Criteria') does not imply similar degrees of biodiversity responses in all ecosystems. Furthermore, we showed how EVD models operationalize this implication further for the purpose of stressor identification and ranking. This approach can, as example, be used for monitoring whether the goals of the (European) Water Framework Directive or the (American) Clean Water Act are reached, and if not, which stressors is or are important (EC, 2000). One of the difficulties in the implementation of these regulations is the full evaluation of variation in stressor levels vis a vis the variation in vulnerabilities, with some pre-set decision criteria which easily lead to Type-1 errors (stressor impacts diagnosed when absent)(Prato et al., 2014). The EVD takes this variation into account and could be used as first tier method for the screening of potential stressor influences for a large amount of locations in a region. Stressor levels lower than the EVD range can be interpreted as being not limitative for the good status of a site; levels higher than the EVD range may imply an affected status; and locations with observations within the EVD range are at risk for having an affected status and should be subject for further assessments. The approach proposed in this paper can also be used in regional planning activities, deciding on the prohibition or allowance of new activities in the region such as a change in land use or the installation of new wastewater treatment techniques. These types of change in activities have impact on different stress-levels at the same time (e.g. nutrients, toxic pressure and pH) and the approach in this paper can be used to model the predicted change in SPO with a changing combination of stress levels for different stressors.

Next to stressor identification and ranking, region-specific boundaries could be set at the safe side of an EVD, e.g. at the 'safe side' (left) of the EVD range or at the stress level at which a selected minimal percentage of the locations exceeds the threshold. Such region-specific boundaries could be applied in risk management and, when assessing ambient stress via sustainability metrics such as the nitrogen footprint method (Leip et al., 2013) and the chemical footprint method (Zijp et al., 2014). Third, the EVD is an operational approach to integrate location-specific impacts across large scale systems. In the context of the planetary boundaries concept, Steffen et al. (2015) pose that various stressors primarily operate at the regional scale, and thus also cross boundaries at local and regional levels. In those cases aggregation to the planetary level is needed to judge global impacts (Rockström et al., 2009b; Steffen et al., 2015). Application in this context is a feasible next step in EVD-development and use. The EVD approach can be used to explore the fraction of sites that are at risk of transgressing a local ecosystem vulnerability threshold for a suite of stressors, given a selected impact threshold or a chosen 'safe' boundary below that.

6.2 Methods

Stepwise principle. The procedure to derive stressor-specific EVDs is visualized in Fig. 6.1 for a hypothetical case with five reference sites.

Data selection and management. Data on the occurrence of the fish species and co-occurring environmental conditions, further referred to as stressors, were available for 1,826 sites sampled between 2000 and 2007 across the state of Ohio, USA (Kapo et al., 2014; Pilière et al., 2014). When multiple measurements were available per location the median value was used. The ecosystem vulnerability distribution approach requires a subset of sites for which the vulnerability distribution is derived as 'point of departure' (Stoddard et al., 2006). In this study, the EVD-derivation and use is illustrated by starting with a subset of 18 sites in stable ecological conditions, here defined as reference sites that are ecologically in balance according to rules of food web allometry (Mulder et al., 2012), a relative novel approach for this purpose in ecology (Mulder et al., 2012) [and references therein]. The resulting selection of 18 reference sites was spatially distributed over the different

Midwestern ecoregions, and was considered to represent the variety of natural conditions in the study area. The reference sites were inhabited by fish communities with an average species richness of 32 (ranging from 12 to 46 species, with one top-predatory species). SDMs were derived for the 84 fish species that were observed at one or more of the reference sites and also at more than twenty of all the sites in the database. Twenty sites is a criterion for the minimum number of data points needed to derive species distribution models (SDMs) (Papeş and Gaubert, 2007; Pearson et al., 2007; Stockwell and Peterson, 2002).

The stressors studied were pH, concentrations of N (total Nitrogen, TN) and P (total Phosphorus, TP), hardness, conductivity, toxic pressure, drainage area and reduction of the Qualitative Habitat Evaluation Index (QHEI). They were selected based on previous studies on these data (De Zwart et al., 2006; Kapo et al., 2014; Mulder et al., 2012; Pilière et al., 2014) while exhibiting low multicollinearity: Variance Inflation Factors (VIF) were below the threshold of 5, transgression of which would imply problems with interpretation (Schipper et al., 2014). The distribution characteristics of the species richness and the stressors are shown in Table S6.1 (Annex, §6.6). Although some stressor data sets contain a few outliers (e.g. 1.9 for pH, or 75 mg/l for TP), the stressors are distributed as expected within a region characterized by a variety of natural background conditions and forms and intensities of human activities. As expected, the reference sites are characterized by relatively low or negligible environmental pressures (Table S6.2). Notably, two reference sites have a QHEI lower than the median for all sites (values: 56 and 59, while QHEImax is 100). Apparently, these assemblages are ecologically in balance while the physical habitat scores is less than optimal according to the QHEI index.

Species Distribution Models. SDMs are quantitative models that combine observations of species occurrence with observations of potential stressors. The shape of the models is $Y= i + c_1S_1 + c_1{'} S_1{}^2 + c_2S_2+c_2{'}S_2{}^2 + c_nS_n+c_n{'}S_n{}^2$, with Y=probability of occurrence, Sn= values for stressors 1-n, $c_{1\text{-}n}$ and $c'_{1\text{-}n}$ are the estimated SDM-coefficients for linear and quadratic terms, respectively and i is an intercept. Ohio fish SDMs were derived following the procedure used by Schipper et al. (2014). Stressor data were standardized prior to model derivation. The SDM models included quadratic terms, to allow for optima/minima. The SDMs were fitted to the data using multiple logistic regressions with a logit link and binomial error distribution with the MuMin package in R (Bartón, 2015). Resulting species-specific models were ranked according to the Akaike's Information Criterion (AIC). Per species, the weighted average regression coefficients were calculated with the MuMin package, based on the set of models within two AIC units from the model with the lowest AIC. Finally, accuracy of the resulting SDMs was evaluated based on the Area Under the Curve (AUC) of the receiver-operating characteristics (ROC) plots using the R package PresenceAbsence (Freeman, 2012).

Reference fish assemblages. Per reference site, the probability of occurrence of each of the fish species observed at the location was estimated with their respective SDM, based on the location-specific stressor information. The basis for assemblage-level stressor-response analyses is subsequently laid by the stacked probability of occurrence (SPO): summing the results of the SDMs (sSDMs) for the local fish species per reference site. The accuracy of each sSDM was assessed by comparing the resulting SPOs with the species richness actually observed at the locations.

Stressor response relationships. Stressor-specific response relationships were derived with the sSDM for each reference site, by in silico changing the stress level of one of the stressors within the range of observed values in the region – leaving out extreme outliers – while maintaining the other stressors at their observed location-specific values. This yields relationships between the value of a stressor and the change in SPO, per reference site.

Environmental Vulnerability Distributions. EVDs were derived per stressor by combining the stressor-response relationships with a selected impact criterion, referred to as threshold. Drainage area was not seen here as a stressor, but as a predictor that describes natural variation amongst the reference sites. The threshold (T) would ideally be based on scientifically derived 'tipping points' (Scheffer et al., 2009). When the stressor-response patterns do not show these sudden responses, an arbitrary choice for T can be operationally used to define and use the EVD, e.g. T defined as the stressor-level at which the predicted sSDM change results in a predicted 5% of species lost. The stress level that results in the response defined by a naturally-defined or chosen T is then calculated with the sSDM per reference site. This yields a range of the critical stress levels across the reference sites that define the distribution of vulnerabilities across reference sites, which is assumed to represent the ecosystem vulnerabilities in the region. The overlay of stressor information for the 1,826 sites and the EVD provides insights for stressor identification and ranking.

6.5 Acknowledgements

Funding by the Strategic Research Program (project number S/607020), the SOLUTIONS-project and the department of Sustainability Environment and Health of RIVM is gratefully acknowledged. The SOLUTIONS project is supported by the European Union Seventh Framework Programme (FP7-ENV-2013-two-stage Collaborative project) under grant agreement no. 603437. We thank R. Heijungs, T. van der Heijde, T. Traas and S.D. Dyer for their suggestions to improve the manuscript and D. de Zwart for support in data preparation. Thanks for the collection and management of the data are due to the Ohio Environmental Protection Agency, the Procter & Gamble Company, and Waterborne Environmental.

6.6 Annex

Table S6.1 Distribution characteristics of stressors and species richness of the fish assemblages in the Ohio biomonitoring database (n = 1826).

Predictor	Mean	StDev	Min	5%	25%	Median	75%	95%	Max
Conductivity (µS/cm)	661	303	101	272	475	642	783	1116	4116
Drainage area(km^2)	111	112	86	86	87	88	96	198	1931
Hardness (mg $CaCO_3$/l)	268	121	32	102	190	265	334	431	1431
pH (-)	7.43	0.73	1.92	6.13	7.33	7.55	7.76	8.13	8.68
QHEI (-)	62	15	16	35	54	64	73	84	101
TN (mg/l)	0.82	0.89	0.11	0.16	0.41	0.67	0.99	1.88	25.00
Toxic pressure, $msPAF_{EC50}$	0.03	0.06	0.00	0.01	0.01	0.01	0.03	0.11	0.72
TP (mg/l)	0.96	5.79	0.00	0.00	0.03	0.06	0.13	2.23	74.44
Fish assemblage									
Species richness (-)	18	8	1	6	12	17	23	34	50

Table S6.2. Distribution characteristics of stressors and species richness of the fish assemblages for the selected reference locations (n = 18).

Predictor	Mean	StDev	Min	5%	25%	Median	75%	95%	Max
Conductivity (µs/cm)	604	197	270	318	488	590	751	874	1002
Drainage area (km^2)	194	103	88	100	116	158	214	363	449
Hardness (mg $CaCO_3$/l)	234	69	121	130	201	219	278	343	353
pH (-)	7.52	0.46	5.92	7.08	7.41	7.54	7.73	7.99	8.15
QHEI (-)	78	11	56	59	72	79	85	92	93
TN (mg/l)	0.94	0.50	0.16	0.21	0.59	0.94	1.23	1.62	2.09
Toxic pressure, $msPAF_{EC50}$	0.03	0.03	0.01	0.01	0.02	0.02	0.03	0.09	0.10
TP (mg/l)	0.16	0.19	0.00	0.02	0.04	0.10	0.20	0.59	0.65
Fish assemblage									
Species richness(-)	32	12	12	13	25	33	42	46	46

Table S6.3 Coefficients and AUCs of the species distribution models per fish included in this study.
msPAF = mixture toxic pressure (at EC50-level), TN = total Nitrogen, TP = total Phosphorus, Cond = Conductivity, QHEI = Qualitative Habitat Evaluation Index, DA = Drainage Area, NA = Not Applicable.

Species names	Intercept	msPAF	msPAF2	pH	pH2	TN	TN2	TP	TP2	CaCO3	CaCO32	Cond	Cond2	QHEI	QHEI2	DA	DA2	AUC
Ambloplites rupestris	-0,08	-0,17	-0,02	-0,38	-0,07	0,11	-0,06	-1,22	0,08	0,70	-0,15	-0,86	0,12	0,65	-0,03	1,70	-0,13	0,78
Ameiurus natalis	0,26	0,31	-0,07	-0,14	-0,03	0,06	NA	0,02	NA	-0,23	0,00	0,38	-0,04	-0,02	0,08	0,33	-0,46	0,61
Ammocrypta pellucida	-28,28	1,66	-0,78	1,10	0,26	-2,49	-2,43	-166,68	-2476,25	1,49	0,02	-2,20	-0,38	0,09	-0,68	2,20	-0,13	0,55
Aplodinotus grunniens	-5,49	0,14	NA	-0,51	-0,12	-1,11	0,21	-59,00	-300,20	-0,25	-0,64	-0,11	NA	0,92	0,03	1,83	-0,10	0,93
Campostoma anomalum	1,51	-0,30	-0,03	0,31	0,00	-0,36	0,08	-0,02	NA	0,74	-0,11	-0,72	0,11	0,77	0,12	-1,20	0,07	0,75
Carpiodes carpio	-2,61	0,69	-0,21	-0,20	NA	-0,64	-0,96	-17,82	-184,41	-0,73	-1,14	-0,45	NA	1,03	-0,07	0,98	-0,07	0,91
Carpiodes cyprinus	-2,49	0,27	-0,13	-0,03	NA	0,07	NA	-0,62	-0,35	0,88	-0,58	-1,12	0,35	0,33	0,04	2,52	-0,14	0,88
Carpiodes velifer	-5,91	-0,12	NA	-0,36	0,03	-1,75	-0,37	-55,25	-346,22	-0,09	-0,93	-0,26	NA	1,23	-0,08	0,84	-0,04	0,93
Catostomus commersonii	2,07	-0,41	0,00	0,35	0,04	0,72	-0,03	0,11	0,04	0,10	0,03	-0,06	NA	0,20	-0,15	-1,73	0,09	0,75
Cottus bairdii	-2,11	-0,67	-0,13	0,09	NA	0,14	-0,03	-5,78	-1,56	0,83	-0,06	-0,92	0,08	0,75	-0,06	-1,50	0,08	0,76
Cyprinella spiloptera	0,48	0,26	-0,06	-0,22	-0,06	-0,07	-0,08	-1,16	0,08	0,25	-0,12	-0,30	0,11	0,25	0,12	9,41	-0,57	0,84
Cyprinella whipplei	-4,24	-0,29	NA	-0,38	-0,50	-2,46	0,20	-10,48	-215,79	-0,78	-2,07	-1,24	0,70	4,03	-1,46	4,04	-0,97	0,98
Cyprinus carpio	-0,77	0,15	-0,04	0,09	-0,03	0,52	-0,02	-1,63	0,11	-0,05	NA	0,37	-0,09	-0,10	0,10	3,94	-0,23	0,81
Dorosoma cepedianum	-1,71	0,37	-0,10	-0,47	-0,12	0,65	-0,27	-1,72	0,14	-0,10	NA	-0,02	NA	0,32	0,00	2,49	-0,14	0,83
Erimystax dissimilis	-49,45	-2,16	-4,53	-1,39	-0,49	-1,11	-8,47	-865,58	-3180,85	3,74	-2,28	-3,75	0,91	4,34	-1,10	0,66	-0,05	0,98
Erimystax x.punctatus	-6,05	1,65	-0,66	0,50	-1,62	-1,98	0,07	-3,37	NA	0,20	NA	0,07	NA	1,14	0,52	1,20	-0,06	0,95
Esox americanus	-1,15	0,20	-0,05	-0,42	-0,07	0,76	-0,45	0,42	-0,03	-0,38	0,06	0,03	-0,06	-0,20	-0,08	1,56	-1,73	0,68
Esox lucius	-14,81	0,71	-1,72	0,39	-0,97	1,49	-1,16	-228,58	-923,08	0,17	NA	0,17	NA	-0,73	-0,33	3,22	-0,59	0,90
Etheostoma blennioides	-0,08	-0,11	-0,04	-0,35	-0,09	0,17	-0,03	-2,62	0,17	0,63	-0,16	-0,60	0,09	0,72	-0,01	0,36	-0,04	0,75
Etheostoma caeruleum	-0,16	-0,06	NA	0,24	0,04	0,00	NA	-3,00	0,21	0,97	-0,20	-0,91	0,15	1,00	0,04	-0,94	0,04	0,79
Etheostoma camurum	-35,87	0,76	-0,79	-0,19	NA	-1,78	-2,09	-469,32	-1826,28	3,09	-1,01	-3,40	0,83	2,73	-0,30	0,92	-0,07	0,96
Etheostoma flabellare	0,01	-0,09	-0,03	-0,43	-0,11	-0,51	0,02	-1,18	0,09	0,33	-0,02	-0,93	0,11	0,52	-0,18	-1,62	0,09	0,79
Etheostoma nigrum	1,22	0,22	-0,14	-0,51	-0,14	-0,26	0,01	-0,35	0,03	0,05	NA	0,05	NA	0,09	-0,06	-1,36	0,07	0,70
Etheostoma tippencanoe	-7,40	0,32	-1,45	-0,40	0,08	-3,12	0,12	-4,10	NA	4,06	-1,94	-2,74	0,65	1,39	-0,35	6,58	-2,26	0,98
Etheostoma variatum	-4,28	-0,02	NA	-0,30	-0,03	-1,99	0,14	-17,96	-179,44	0,35	NA	-0,60	-0,16	1,87	0,39	0,37	-0,03	0,92
Etheostoma zonale	-1,71	0,15	-0,04	-0,16	-0,05	0,00	NA	-0,94	-0,19	0,91	-0,23	-1,11	0,20	0,94	0,12	0,82	-0,05	0,80
Fundulus notatus	-2,29	0,57	-0,06	-0,38	-0,22	0,33	-0,01	-4,16	0,34	1,63	-0,54	-0,49	0,10	-0,62	0,01	1,88	-2,00	0,82
Hybopsis amblops	-4,56	-0,65	-0,31	-0,03	NA	-0,63	-0,18	-3,12	-30,32	1,08	-0,21	-1,94	-0,26	1,44	-0,20	0,72	-0,05	0,87
Hypentelium nigricans	0,42	-0,07	-0,04	0,07	-0,03	-0,55	0,02	-0,68	0,05	0,08	NA	-0,37	0,06	0,99	-0,05	4,29	-0,27	0,84
Ictalurus punctatus	-1,71	0,90	-0,21	-0,32	-0,12	0,51	-0,20	-0,33	NA	0,33	-0,15	-0,28	NA	0,45	-0,01	4,01	-0,23	0,91
Ictiobus bubalus	-4,17	0,45	-0,11	-0,93	-0,43	-1,20	-0,06	-28,00	-266,78	-0,52	-0,65	-0,14	NA	1,82	-0,27	0,74	-0,03	0,96
Labidesthes sicculus	-2,98	0,58	-0,17	-0,15	NA	0,05	NA	-0,47	-0,62	0,46	-0,65	-1,42	0,30	0,46	-0,02	0,60	-0,04	0,78
Lampetra aepyptera	-3,73	-0,21	NA	-0,31	-0,02	-1,08	-0,03	-1,43	0,12	-0,10	-0,34	-1,62	0,27	0,17	-0,19	-0,37	NA	0,87
Lepisosteus osseus	-4,44	-0,05	NA	0,08	NA	-1,35	0,07	-6,14	-89,24	-0,31	-0,55	-0,12	NA	1,30	0,20	0,96	-0,04	0,90
Lepomis cyanellus	1,39	0,29	-0,09	-0,22	-0,04	0,60	-0,02	0,02	NA	-0,05	NA	0,22	-0,03	0,19	-0,06	-0,37	0,02	0,65
Lepomis gibbosus	-3,31	0,05	NA	0,33	0,03	0,66	-0,18	-9,25	0,73	-1,84	0,07	1,31	-0,27	-0,03	NA	0,57	-0,19	0,78
Lepomis gulosus	-2,68	0,43	-0,31	-0,16	0,02	-0,23	-0,10	0,09	NA	-1,46	-0,32	0,81	-0,07	-0,31	-0,09	2,29	-2,69	0,76
Lepomis humilis	-3,27	0,72	-0,08	-0,43	-0,14	0,99	-0,16	-3,90	-7,10	0,95	-0,35	-0,47	0,14	-0,12	-0,11	0,88	-0,05	0,81
Lepomis macrochirus	0,81	0,14	-0,06	0,05	NA	0,09	0,00	0,03	NA	-0,11	NA	0,11	NA	0,39	0,08	0,70	-0,05	0,64

Lepomis megalotis	-0,76	0,43	-0,07	-0,59	-0,17	-0,11	0,01	-1,08	-0,03	0,78	-0,14	-0,83	0,16	0,19	-0,05	1,17	-0,24	0,72
Lepomis microlophus	-3,51	-0,22	NA	-0,06	NA	0,80	-0,77	0,74	-0,20	-1,20	-0,53	-0,52	0,08	0,06	NA	0,68	-0,03	0,79
Luxilus chrysocephalus	0,48	0,30	-0,07	-0,02	-0,04	-0,33	0,02	-0,26	0,03	1,07	-0,25	-1,23	0,21	0,53	-0,09	-0,36	0,02	0,74
Luxilus cornutus	-1,87	0,29	-0,16	0,37	0,06	0,16	-0,01	-1,83	0,14	-1,08	0,17	1,18	-0,27	0,07	0,08	-0,57	0,03	0,72
Lythrurus fasciolaris	-0,15	-0,07	NA	-0,49	-0,10	0,02	NA	3,14	-16,82	1,50	-0,45	-2,17	0,35	0,34	-0,13	1,52	-8,86	0,79
Lythrurus umbratilis	-1,42	0,76	-0,22	-0,51	-0,19	0,16	0,00	-1,44	-5,85	0,46	0,01	0,18	-0,14	-0,59	-0,12	3,48	-3,49	0,75
Micropterus dolomieux	0,24	-0,12	-0,04	0,08	NA	-0,04	NA	-1,55	0,07	0,79	-0,16	-0,91	0,19	0,71	-0,07	8,26	-0,54	0,87
Micropterus punctulatus	-2,23	0,60	-0,29	-1,06	-0,21	-0,73	0,03	-0,27	-0,03	-0,12	-0,22	-0,28	0,09	0,82	-0,22	1,05	-0,06	0,84
Micropterus salmoides	-0,13	0,02	-0,04	0,02	NA	0,41	-0,11	-0,28	0,03	-0,29	0,03	0,24	-0,03	0,14	0,00	0,04	NA	0,61
Minytrema melanops	-1,88	0,22	-0,22	-0,43	-0,09	0,57	-0,60	0,00	NA	0,07	NA	0,10	NA	-0,02	-0,16	2,66	-0,85	0,79
Morone chrysops	-16,61	0,43	-0,17	-0,30	-0,05	-0,25	NA	-154,06	-978,22	-0,81	-0,90	-0,75	-0,92	1,75	-0,31	0,39	-0,02	0,86
Moxostoma anisurum	-2,31	0,70	-0,26	-0,67	-0,20	-0,74	0,03	0,08	-2,97	0,07	-0,11	0,07	NA	0,65	-0,17	3,22	-0,19	0,91
Moxostoma breviceps	-3,15	-0,78	0,08	-1,35	-0,23	-0,28	NA	-8,83	-162,14	-0,40	-2,16	-1,54	0,56	3,49	-0,82	0,96	-0,05	0,97
Moxostoma carinatum	-9,36	1,11	-2,10	0,56	0,10	-0,09	NA	-77,62	-639,43	0,35	-0,06	0,36	NA	1,88	-0,51	1,15	-0,05	0,92
Moxostoma duquesnei	-1,44	0,00	NA	-0,04	NA	-1,00	0,04	3,97	-16,85	1,15	-0,57	-1,76	0,49	1,14	0,12	3,34	-0,72	0,90
Moxostoma erythrurum	-0,27	0,01	-0,04	-0,68	-0,11	-0,25	0,01	-0,90	0,07	0,69	-0,29	-1,02	0,20	0,61	0,00	5,53	-0,34	0,87
Moxostoma macrolepidotum	-22,51	1,00	-0,21	0,40	0,09	-1,58	0,36	-320,48	-1310,36	0,32	-0,11	0,21	NA	0,32	-0,24	2,35	-0,29	0,90
Nocomis biguttatus	-3,20	1,16	-0,41	0,54	0,04	0,63	-0,87	-3,62	NA	1,59	-0,35	-0,94	0,14	0,33	0,14	-0,16	-6,03	0,78
Nocomis micropogon	-4,37	-0,23	-0,02	0,38	0,10	-0,21	-0,44	-5,81	-7,08	-0,98	-0,43	0,27	NA	1,57	0,00	1,65	-0,49	0,89
Notemigonus crysoleucas	-2,27	0,25	-0,05	0,09	NA	0,56	-0,14	0,00	NA	-0,71	0,09	0,58	-0,06	-0,55	-0,12	0,20	-0,48	0,70
Notropis atherinoides	-2,07	1,30	-1,03	-0,41	-0,53	-0,11	NA	3,51	-7,65	-0,28	-0,39	-0,06	-0,36	0,67	0,11	0,91	-0,05	0,86
Notropis buccatus	-0,41	0,10	-0,07	-0,35	-0,08	-0,73	0,11	-0,56	0,04	0,22	-0,08	-0,41	0,09	0,05	-0,07	-1,22	-0,13	0,70
Notropis photogenis	-1,01	0,05	NA	-0,25	-0,08	-0,37	0,02	3,09	-18,61	1,09	-0,26	-1,31	0,24	1,22	0,04	0,62	-0,06	0,83
Notropis rubellus	-2,72	0,42	-0,14	-0,34	-0,11	-0,75	0,03	-2,65	-3,64	0,32	-0,19	-0,21	0,17	1,11	0,19	1,09	-0,23	0,83
Notropis stramineus	-0,81	0,23	-0,09	-0,09	-0,02	-0,18	0,01	-0,64	0,01	0,43	-0,15	-0,38	0,11	0,45	0,07	1,49	-0,09	0,74
Notropis volucellus	-3,21	0,54	-0,20	-0,20	-0,05	-0,21	NA	-1,03	-4,85	0,06	NA	-0,63	-0,11	1,12	0,04	0,79	-0,05	0,85
Noturus flavus	-2,25	0,16	-0,04	-0,08	-0,07	0,10	NA	-0,58	-0,01	0,55	-0,28	-0,63	0,11	1,25	0,04	1,96	-0,51	0,86
Noturus miurus	-3,78	-0,27	-0,29	0,10	NA	-0,88	-0,07	-1,27	-7,61	1,78	-0,65	-3,38	0,43	0,47	-0,27	3,72	-1,78	0,88
Perca flavescens	-4,27	-0,52	-0,11	0,16	-0,07	1,59	-1,01	-10,09	0,75	-0,89	0,02	0,09	NA	-0,23	0,03	1,89	-0,35	0,80
Percina caprodes	-1,00	0,11	-0,05	-0,43	-0,12	0,41	-0,09	-1,26	0,10	0,68	-0,14	-0,81	0,14	0,76	0,01	2,42	-0,15	0,83
Percina maculata	-0,84	0,51	-0,15	-0,23	-0,09	0,16	-0,09	-0,64	-0,01	0,11	NA	-0,22	-0,02	-0,09	-0,16	1,50	-0,64	0,67
Percina phoxocephala	-4,73	-0,09	NA	-1,03	-0,51	-1,05	0,14	-32,30	-266,55	-0,32	NA	0,03	NA	2,10	-0,34	0,57	-0,03	0,95
Percina sciera	-3,98	1,93	-0,71	-0,21	-2,50	-3,53	-2,52	-15,94	-190,39	-0,17	NA	-0,05	NA	0,54	-0,32	2,71	-0,53	0,92
Percopsis omiscomaycus	-2,89	0,33	-0,11	-0,28	-0,09	-0,44	-0,07	-5,20	-58,05	-0,62	0,06	-0,08	NA	0,12	-0,27	0,66	-0,11	0,72
Phenacobius mirabilis	-2,19	0,72	-0,12	0,41	-0,65	-0,14	0,01	-0,06	-6,16	0,50	-0,14	0,45	-0,03	0,09	NA	0,68	-0,04	0,75
Pimephales notatus	2,01	0,06	-0,06	0,13	0,00	0,14	NA	-0,60	0,05	0,16	-0,07	-0,07	NA	0,29	0,11	0,56	-0,05	0,69
Pimephales promelas	-1,75	0,37	-0,06	0,43	0,07	0,34	-0,01	-0,23	0,01	0,33	-0,03	0,19	-0,02	-0,47	0,00	-0,54	0,04	0,73
Pimephales vigilax	-5,87	0,04	NA	-1,06	-0,41	0,40	-1,27	-5,34	-97,65	-0,78	-0,38	-0,13	NA	1,61	-0,33	1,25	-0,07	0,96
Pomoxis annularis	-1,83	0,79	-0,24	-0,33	-0,13	1,02	-0,47	0,25	-0,02	0,02	-0,19	0,03	NA	0,25	-0,28	0,48	-0,02	0,75
Pomoxis nigromaculatus	-2,62	0,64	-0,50	0,31	-0,02	0,29	-0,02	-0,20	NA	-1,08	-1,14	0,18	NA	0,49	0,01	0,47	-0,02	0,81
Pylodictis olivaris	-4,24	0,86	-0,34	-0,68	-0,17	-3,09	-1,40	-49,85	-365,61	-0,17	-0,82	-0,87	0,42	1,28	-0,31	2,30	-0,13	0,96
Rhinichthys obtusus	-0,68	-0,56	0,04	0,60	0,09	-0,16	0,01	0,20	0,01	0,12	0,03	-0,24	0,05	0,39	-0,13	-5,13	0,32	0,79
Sander canadensis	-3,10	0,90	-0,53	-0,56	-0,45	-2,97	0,22	-28,39	-247,71	0,25	-0,49	-0,10	-0,57	1,16	-0,26	1,08	-0,08	0,93
Semotilus atromaculatus	1,95	-0,39	-0,01	0,23	0,03	0,09	0,15	0,30	0,02	0,58	0,03	-0,45	0,03	0,06	-0,11	-3,93	0,23	0,89
Umbra limi	-3,15	-0,40	-0,07	-0,28	-0,27	0,37	-0,04	-1,58	-19,07	-0,85	0,15	0,51	-0,28	-0,36	-0,04	-4,62	-14,09	0,76

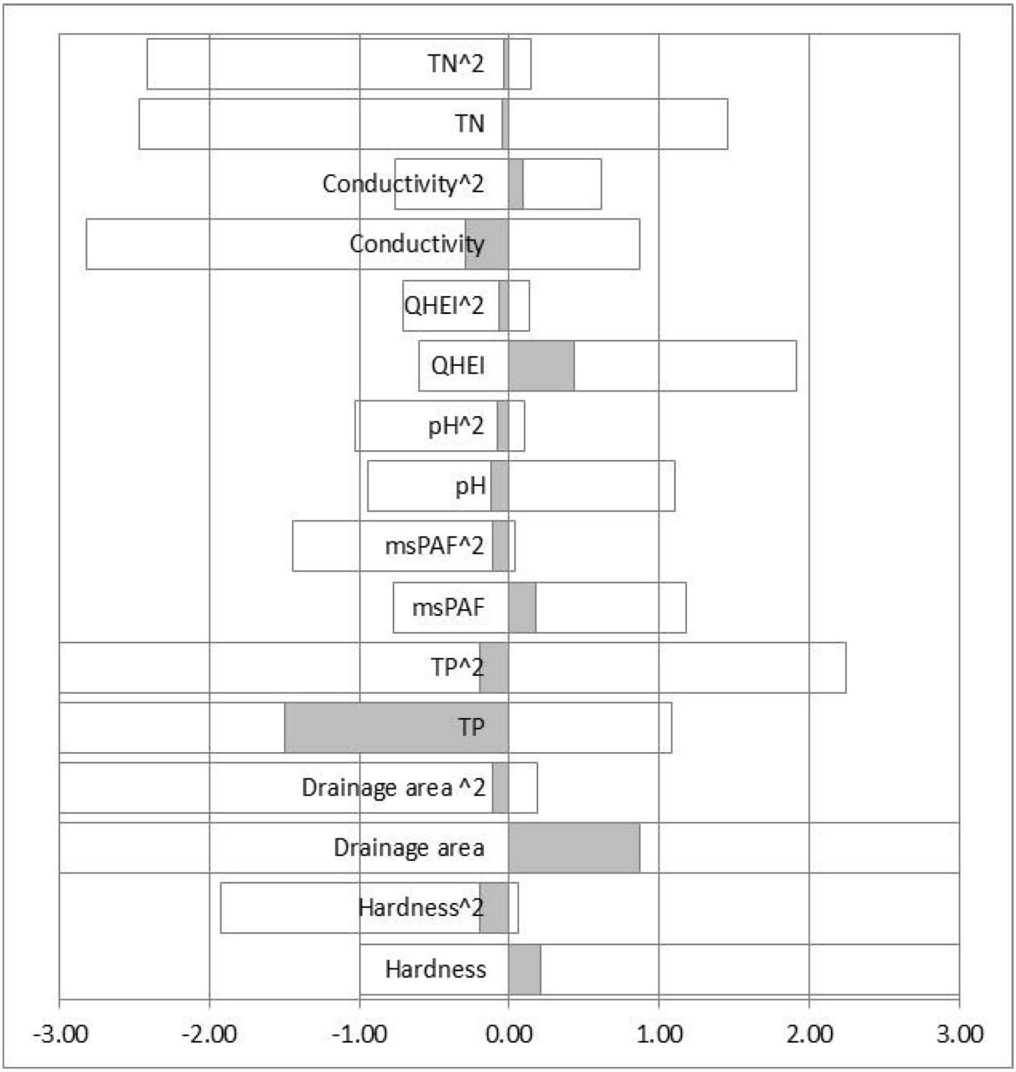


Fig S6.1 Across-species variability of the standardized coefficients c and c' (linear and quadratic, respectively) for the different stressors of the SDMs for the 84 fish species. The grey bar is the median- and the bars on the left and right respectively the 5th and the 95th percentile of the coefficients across the species set. Values for the 5th and 95th percentiles that exceeded the figure bounds where cut off.

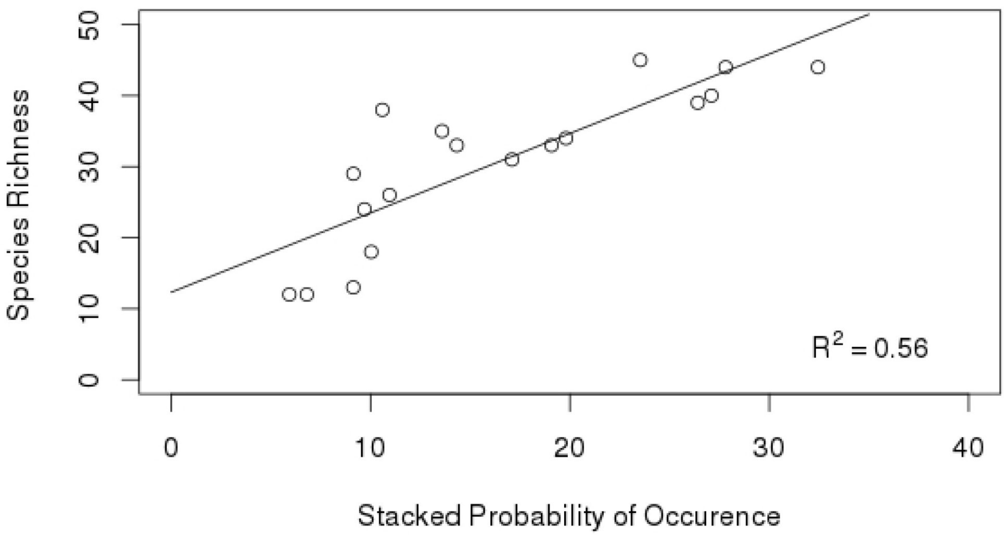


Figure S6.2 Significant association between measured species richness (y-axis) and the modelled stacked probability of occurrence (SPO, x-axis) for the 18 reference sites, $R^2 = 0.56$, $p < 0.01$. SPO is always lower than the species richness, as a probability is a number between 0 and 1 and the measurement of occurrence is binomial.

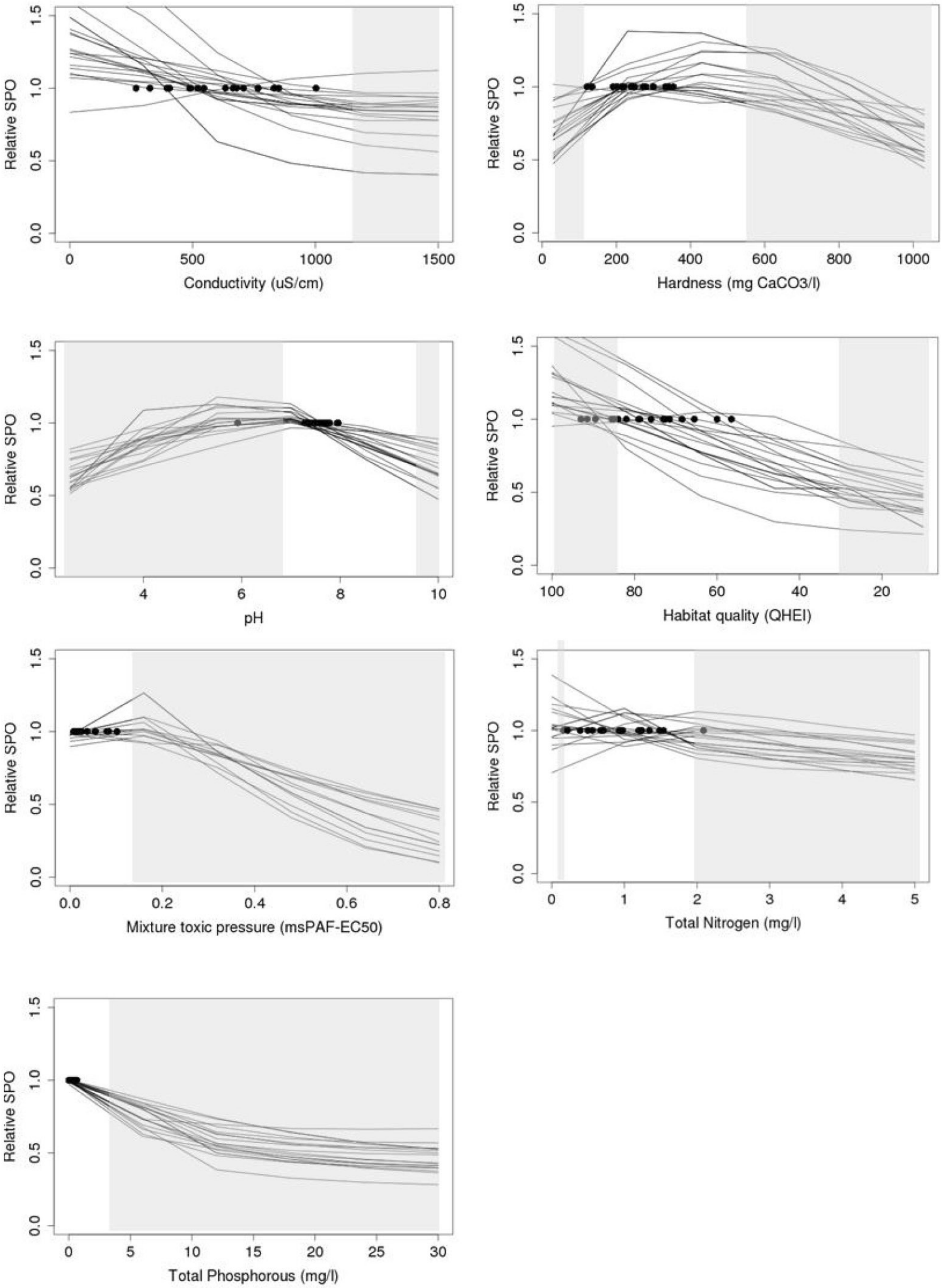


Fig S6.3. Dose response curves for Total N, pH, Conductivity, QHEI, Hardness, Toxic Pressure (msPAF) and Total P, with at the y-axis the relative SPO, which is the stacked probability of occurrence compared to the SPO at observed field values. For the SPO in this figure only the SDMs with an AUC higher than 0.7 are used.

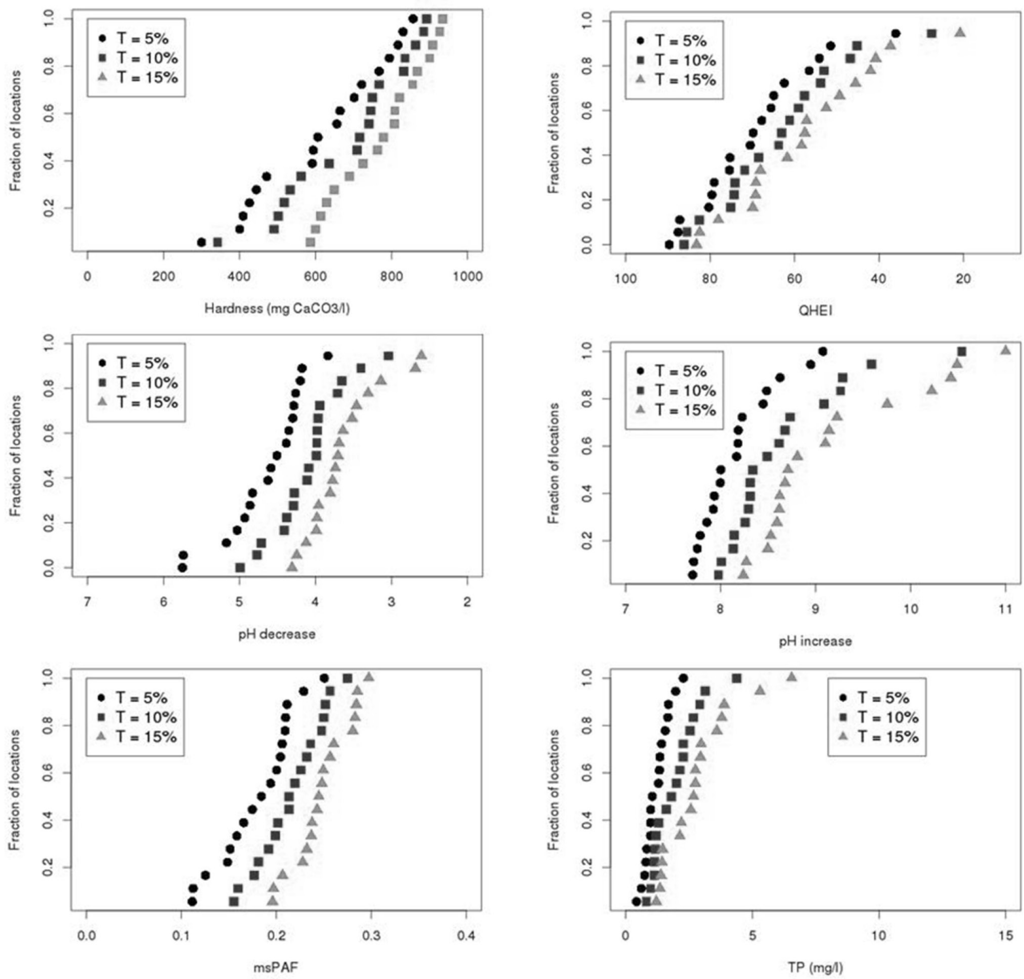


Fig. S6.4. Example EVDs with different choices for a selected Threshold T on the stressor-response models of Panel A in Figure 1 (main text). Main text results for T=5% SPO decrease compared with SPO-reductions of 10% and 15% on the stressor-response models.

7 Synthesis

7.1 Increase the utility of sustainability assessments

Have you read the United Nations Sustainable Development Goals (UN, 2015a)? The list of challenges we currently face as humanity is impressive: achieving a social humane foundation for everybody within the carrying capacity of our planet. The challenges are visually summarized in the conceptual model of Raworth (2012) (Fig. 1.1), acknowledging the interconnectedness of environmental and social issues. Meeting the challenges requires well-considered choices to forward sustainable development. Those are made daily, in different situations, by different people at different places on different spatial scales and temporal horizons at different times. Sustainability assessments (SA) are meant to support these choices. However, in the introduction of this thesis we argued that the supporting power of sustainability assessments is hampered by:

1. the tendency to focus assessments on problems instead of possible solutions (Clark and Dickson, 2003; U.S. NAS, 2009);
2. SA-method selections that do not fit the decision context and the view of the involved stakeholders on sustainability (Gasparatos and Scolobig, 2012; Waas et al., 2014); and
3. the struggle with environmental boundaries: how do we derive boundaries, especially for those with apparently distributed stressor and vulnerability levels, and how do we use them in governance to forward sustainable development (Pope et al., 2004; Sala et al., 2013)?

The Chapters of this thesis are concerned with these three limitations of common, current SA's. The thesis describes new approaches for SA-process improvement, new approaches for addressing boundary-related discussions and case studies involving both process and boundary aspects. Note that, the contributions to the derivation and use of boundaries and the case studies in this thesis focus on the environmental part of sustainable development, while the procedures and lessons learnt are applicable for sustainable development in its broad definition (see introduction).

Sustainability assessments are typically required in situations that are characterized by multiple interconnected environmental and social issues, multiple perspectives of different stakeholders on the problem and trade-offs between conflicting goals (Kates et al., 2001). As a consequence, typical SA cases concern situations in which there is no single optimal sustainable solution that is easily accepted by all stakeholders (Table 2.1, Chapter 2). In order for a sustainability assessment (SA) to contribute to decisions in such 'wicked' situations, it has been proposed that they should be embedded in a procedure with the following features (Table 2.2, Chapter 2):

- Participative: engaging stakeholders in the process;
- Iterative: not finished after a decision. The decision phase is only one step in the search on how to deal with the problem. The process keeps on evaluating ways to improve the sustainability status of the wicked situation, i.e., continuous sustainable development;
- Solution-focused: possible solution scenarios are discussed upfront in the assessment process, also when there is no consensus on the problem definition.
- Innovative: solutions are sought 'out of the box', e.g. solutions might be found that require reconsideration of existing policy standards and boundaries.
- Clarity about its definition of sustainable development: the translation of sustainable development into the multiple metrics by which it becomes possible to assess the benefits and costs of the solution scenarios is done transparently.

An approach that addresses all these criteria was designed (see Fig. 7.1 and chapter 2). This was done by combining existing frameworks from the realms of risk assessment (U.S. NAS, 1983), risk governance

(IRGC, 2008), adaptive management (Linkov et al., 2006; Robinson and Levy, 2011), sustainability assessments (Gorman et al., 2012; Sala et al., 2013; Zijp et al., 2015), and - especially - a solution-focused paradigm superimposed on the combined framework elements (U.S. NAS, 2009). The innovative parts of the framework introduced in Chapters 2, 3 and 4 are the exploration of alternative management solutions upfront in the risk- or sustainability assessment process ánd the explicit step of SA-methods selection (setting the rules for the sustainability assessment and identifying the pertinent metrics to represent and evaluate the problem and its optional solutions). The planned benefits of the solution-focused sustainability assessment (SfSA, Ch. 2) framework, combined with the specifications of the method selection using the sustainability assessment identification key (SAIK, Ch. 3 and 4), are that assessors become actively involved in identifying the key societal questions as well as the full width of potential solutions to minimize risk and optimize sustainable development. Consequently, the assessment results might be closer to (or better applicable for) management decision support (U.S. NAS, 2009).

The SFSA-SAIK framework describes how a sustainability assessment is part of a sequence of steps, with possible iterations between all steps and that the sustainability assessment does not take place before some other important activities have taken place: the pre-assessment, the search for solution scenarios and the specifications for the sustainability assessment method selection. This thesis contributed to the design and testing of these process steps, including the sustainability assessment step, with new procedures, metrics and case studies (Fig. 7.1). In the following sections (7.2-7.5) these are further explored in relation to sustainability assessments. It is noted that the illustration of SfSA with the case study on management of contaminated sediments covers all steps of the SfSA procedure, and that the chemical footprint in chapter 5 and Ecosystem Vulnerability Distributions provided in chapter 6 are also used for evaluation of realisation of the planned sustainable development improvements (evaluation step of SfSA). However, the focus of this synthesis is on the first four steps of SfSA towards and including the SA.

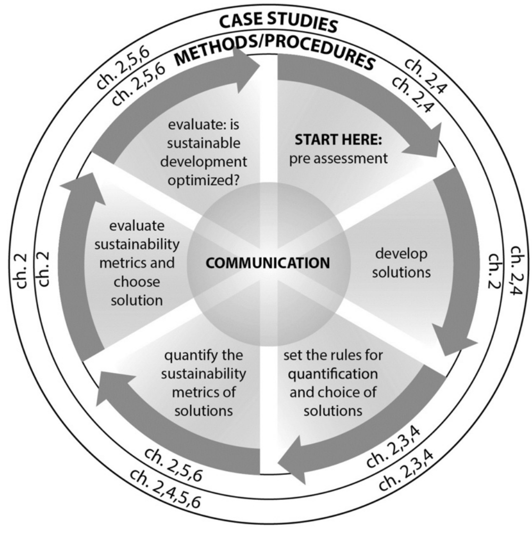


Fig. 7.1 The design of the Solution-focused Sustainability Assessment framework and its association with the different chapters in this thesis.

7.2 The pre-assessment

The pre-assessment is meant to identify the specifications of the sustainability problem and gather information on the context. The information types gathered vary widely: scientific information from previous research, regulatory information on for example relevant policy standards, the underlying first principles of those standards and stakeholders identification. This information is then used to discuss the design of the assessment process: which activities with which iterations are foreseen within which time-period, and: who should be involved in the different stages of the SfSA framework. The operationalization of SfSA is case specific. For example, designing a national policy to operationalize the dietary fish consumption advice of a health council (Hollander et al., in prep) will likely involve more sensitive stakeholder dialogues to design a stakeholder-supported sustainability assessment of solution scenarios, then a drinking water companies' choice for one out of four alternative techniques to execute one raw water clean-up step during a drinking water production process (Van der Laan et al., 2016). The former requires wide societal involvement while the latter involves mainly drinking water management stakeholders.

During the pre-assessment, the contours of the requirements for the sustainability assessment are already shaped. The information gathered provides insight in which parts of the problem are and are not yet characterized or quantified in literature.

Stakeholders should already be involved in this stage, for example to define the problem and its context from a suite of viewpoints. The protocol provided in Chapter 4 (§4,7,3) is designed to support this pre-assessment. Involving stakeholders in such early stage of a project is currently rare. The first obstacle is not to get stakeholders involved, but the fact that 'problem owners' seem to have the tendency to first make up their own mind on the problem, its context, the possible solutions and definition of sustainable development, before consulting the view of other stakeholders. The common approach is 'inside the known box'. This tendency was for example seen in the case study on sediment management (Chapter 2) and the Resource and Energy Factory (chapter 4). Although understandable, from the perspective of sustainable development it seems more valuable to start the (re)definition of the problem and the search for optional solutions together, because sustainable solutions require a transdisciplinary view, which is only gained when collating and exploring insights and views from different perspectives (Gorman et al., 2012). In literature this is called: deriving a trans-disciplinary conceptual model (Pooley et al., 2013; Sala et al., 2012b), or finding the trading zone for solutions. Defining a trans-disciplinary view on the problem and on sustainable development is easier when the different disciplines have not yet specified their own view, but dare to define their own view together with others (Gorman et al., 2012). In general the 'solution-space' will be larger when it is not implicitly blocked by unquestioned "beliefs" and when the project is a co-creation from the start.

Most of the complex sustainable development problems require involvement from the start of researchers and/or the experts that will later (step 4) perform the sustainability assessment (UN, 2015b). Involvement creates a broad understanding of the complex multi-disciplinary context, which in turn allows the researchers and experts to design a sustainability assessment that might fit the decision context, including possible trade-offs, and which is supported by the stakeholders. The latter is, next to scientific findings, key for implementation of a preferred sustainable-development solution. It is the implementation of appropriate solutions that forwards sustainable development, not solely the quantitative findings of a SA per se. Furthermore, scientist can contribute by sharing their view on the situation and propose solutions others might not be thinking of. Research designed and performed directly in the context of an application is a key feature of sustainability science according to, amongst others, Kates et al. (2001).

7.3 Alternative solution scenarios

The next step, after the pre-assessment, is to find solution scenarios for the problem(s). Although this might seem evident, in practice the first steps after pre-assessment are efforts to define one problem definition and to perform research on the problem: Is there a problem? What are the risks of transgressing a boundary? Which stakeholder's view is the right one? This might be the reason most published sustainability assessments appear to be problem focused instead of the by sustainability science pursued and by society requested focus on solutions (Clark and Dickson, 2003; Kates et al., 2001; Sala et al., 2012b; UN, 2015a).

Here a different strategy is proposed: acknowledge the potential wide variability of problem definitions, but leave these different problem definitions for what they are for a while, and start a process in which it is possible to define a wide and relevant set of solution scenarios for the future. With other words: shift the focus from the problem to the solution(s). This strategy has three benefits: 1) the focus on solution alternatives provides a new view on the present situation and on the perspectives of the different stakeholders; 2) subsequent research is more tuned to management and the decision procedure (U.S. NAS, 2009); it focuses on sustainable development, instead of solely on quantitatively specifying the risks of the current situation; and 3) instead of focusing on differences in problem definition, it seeks for the shared solution space (Gorman et al., 2012). Stakeholders with different problem perceptions can agree on the level of solutions, but for different reasons. For example, both climate-change 'believers' and 'skeptics' can agree on measures to reduce the dependency on fossil fuels for different reasons, i.e. the 'believer' to reduce climate change, the 'sceptic' for political or economic reasons.

Many handbooks and other descriptions are available on how to perform a successful brainstorm exercise in exploring a problem definition and its potential solution space with different stakeholders (Zijp et al., 2016). This activity does not only include gathering as many solution scenarios as possible, but also to decide which of them are to be considered in the sustainability assessment steps and the subsequent decision procedure. That is, the process yields the most promising alternative scenarios.

The brainstorms on solution scenarios should not directly be delineated by existing risk boundaries or policy standards, but by first principles: what is to be achieved (social foundation) and what is to be protected (environmental health)? In this step, it should be acknowledged that the existing boundaries and standards – like the planetary boundaries – consist of a scientific and a normative part (e.g. safety factors) and are often set from a mono-disciplinary perspective, to reach a single goal with a high degree of protection. In order to provide room for solutions to multi-disciplinary problems the normative choices on which the boundaries and standards are based need (at least) be known, while they might be reconsidered when exploring the solution space. That is, as shown in Fig. 1.3, the current boundary for a problem may, when transgressed, not lead to immediate impact, which thus leaves room for evaluation.

It is important that scientists are involved in this step towards exploring potential sustainable solutions. Scientists can drive the transition towards a solution-focused approach. This asks from them, however, a solution-focused mindset, as well as a critical approach to the definition of sustainability assessment projects. It is already the question types that are directed at scientists that are often problem focused and mono disciplinary. For example, three different governmental bodies ask the Dutch Research Institute for Public Health and Environment (RIVM) to report on specific aspects of Dutch food consumption: one asks to monitor the safety of Dutch food consumption and provide policy advice from that to maintain or improve food safety, one asks to monitor the health aspects of the Dutch food consumption, and provide policy advice on measures to increase healthy food consumption, and one asks to monitor the environmental impacts of the Dutch food consumption, and provide policy advices on how to reduce the environmental impacts of the Dutch food consumption pattern. The separate requests resulted in three different monitors (De Valk et al., 2016; Friesema et al., 2014; Van Rossum et al., 2011), which may be utilized in three different, unrelated decision procedures on (policy) measures with regard to food consumption: health, safety and environmental impacts. However, measures with regard to safety have direct consequences for the environmental impact of the food system, e.g. via packaging material use, or via defining the stage where food becomes unsafe, and should be seen as waste. In other words: an improvement of one, may imply a trade-off elsewhere. As an alternative, all three should be considered simultaneously in order to uncover possible trade-offs and pursuit sustainable development. Bridging such widely different disciplines, as recently suggested for the food dossier (Ocké et al., 2017), will provide an overarching and balanced advice for a transition towards a sustainable, healthy and safe food consumption pattern and a research agenda that focusses on how to anticipate on – and counteract potential trade-offs.

To strike the balance between risks and sustainable development can be a highly complex enterprise, is shown in a recent example, which directly involved environmental quality and human health. In 2016, there was societal unrest with regard to potential human health impacts of granulated rubber used on sport fields. The material originates from car tires. The question directed to the RIVM (but a comparable question is directed to the European Chemical Agency (European Commission, 2016b)) was if there is indeed a problem for human health. This was assessed in a human-health risk assessment, resulting in the human health risk assessment outcome that "the health risk of playing sports on synthetic turf pitches with an infill of rubber granulate is virtually negligible" (Oomen and De Groot, 2016). However, from the viewpoint of sustainable development, a pre-assessment of the problem in the first phase of the SfSA would have resulted in a discussion on the broader perspectives on the problem. It would have uncovered, for example, the existence of various liaisons to other dossiers such as on emissions of micro-plastics to the environment and on the essentials of realizing a circular economy (improved resource management). Furthermore, applying the second phase of the SfSA (finding solutions before assessing sustainability) would have resulted in a list of alternative solutions for granulated rubber to be evaluated, which, in turn, would have led both to insights in the risks, as in possible ways to deal with these risks. This extension of the focus from: "Is there a specific risk problem from a mono-disciplinary point of view?" to: "Is the use of rubber granulates on sport fields smart from a sustainable development point of view?" is what is meant with 'Solution-focused'. It requires an investment prior to the assessment, but leads to an increased utility of its outcomes. Scientist can support this transition by taking the initiative, at every opportunity, by questioning the question and exploring creative ways of solving complex problems (Kibwika and Wals, 2008).

7.4 Setting rules for the sustainability assessment

After identifying solution scenarios and selecting some that are implied to potentially contribute to sustainable development, in general the next step is to perform a sustainability assessment to compare the alternatives. The idea is to identify the solution that contributes most to sustainable development. In practice, the choice for methods to perform the assessment is driven by the available expertise (Zijp et al., 2015). This is problematic, because the sustainability assessment needs to be the concrete translation of sustainable development in metrics. It is in general not possible to include all aspects of sustainability in one assessment. Therefore, choices are made with the method selection where to focus on. Because different stakeholders tend to have different views on sustainable development there is chance they will not recognize their view in the focus of the assessment and therefor do not support the results, hampering the use of the assessment results during the decision-making process (ibid.). This is why an expert from the team that will perform the assessment should be involved in the previous phases. This is also why an extra step is proposed between the definition of solutions and the sustainability assessment itself: setting the rules for the assessment.

A method selection that fits the view of stakeholders on sustainable development requires specification of the system boundaries (e.g. spatial scale) of the assessment, the themes to be considered and the type of aggregation (i.e. normalization and weighting) that may be chosen or will be required. Additionally, criteria on the procedural design (e.g. stakeholder involvement) and organisational restrictions of the assessment (e.g. data availability) could be considered. Thus, as a matter of fact, the question articulation does not stop at the pre-assessment (first phase), but it finds an extra part in this phase, when the requirements for the sustainability assessment are delineated (Zijp et al., 2015).

Ideally, this question articulation is derived together with stakeholders and aligned with existing literature on assessments in comparable situations. However, case studies performed with the identification key (a.o. Chapter 4) revealed that problem owners in the first instances commonly want to make up their own mind. This can be dealt with by iterative method selection (first in a workshop with internal stakeholders and with an optional adjustment of the outcomes with another workshop with external stakeholders) and by adding workshop elements that challenges internal stakeholders to think from the perspectives of the external stakeholders (Chapter 4, §4.7.3).

The specifications defined by such explorations can be linked to the available SA methods, for which we defined the sustainability assessment identification key (SAIK, chapter 3). A first step is made in operationalizing the SAIK. A protocol was designed with which the demands for the assessment can be defined together with stakeholders and a tool was made to match those demands with a selection of available methods: www.sustainabilitymethod.com. Both the protocol and the tool focus on the criteria for delineating system boundaries and theme selection and should expand to aggregation and organisational restrictions in the near future. Application of the protocol and tool, amongst others in a case study on the recovery of resources from waste water (Chapter 4), revealed that method selection is an iterative process. The availability of a tool for SA-method selection does not diminish the role of the sustainability expert. Expert knowledge is required to propose a consistent set of themes based on all input of the stakeholders and, when necessary, propose a combination of complementary methods when there is not one exact match. The SAIK operationalizes this third phase of the SfSA framework.

The result of this phase is case specific, depending on the problem, its context, the solutions that can be defined, and the stakeholders involved. Consequently, a set of SA methods that exactly matches the specifications provided will often not be available. In those cases a combination of existing methods can be used, for example, the case of recovery of resources from domestic waste water (Chapter 4) required a combination of life cycle assessment and expert elicitation. Another option is to transparently alter the specifications, which also might be necessary due to other organizational restrictions such as data availability. The transparent alteration concerns that all the important choices made in the translation process are discussed with the problem owner and are reported. This provides clarity on what can be expected of the assessment before it takes place, and explanations on choices in the assessment design when the results are debated during the decision procedure.

Sustainability assessment experts have an important role during this phase. Firstly, based on knowledge and experience in the field of sustainable development they can contribute in the discussion on what is important to include, as one of the stakeholders. Secondly, the specification can lead to a list of selected sustainability themes that partly overlap. The expert can propose a coherent set of themes based on the input of the question articulation. Then, thirdly, when the list of possible (combinations of) methods are derived, e.g. with the tool at www.sustainabilitymethod.com, the expert can propose which combination suits best, given the organizational restrictions such as data availability.

Using the sustainability assessment identification key is an investment at the start of the project in time and effort, as compared to the current processes in many sustainability assessment projects. However, it results in a more transparent method selection, clear expectations on the assessment outcome types, and a method that fits the decision context as good as possible, including the stakeholders' views on sustainable development.

7.5 Selecting and applying the sustainability assessment methods

As a next step, when the rules for the assessment are set and methods are selected, the sustainability assessment can take place. The activities in this phase do not only lead to a set of sustainability scores for the different metrics for the different solution scenarios. It can also lead to: i) an adjusted selection of solution scenarios or adjustments within the scenarios, e.g. changing part of a technology that shows high contributions on the total impact (Zijp and Van der Laan, 2015); and ii) an evaluation of the relation across the partners in a production chain. For example, gathering data on the life cycle impacts of a product requires data from suppliers. The willingness to cooperate in realizing transparency in the production chain is an indicator for the trust between partners and can be interpreted as the potential for collaborative activities in the chain towards sustainable development (Zijp and Van der Laan, 2015).

In practice, the sustainability assessment will often be an iterative process that includes re-checks of activities in the previous steps. For example, a quick scan assessment can be used to narrow down the list of possible solutions considered before performing a thorough assessment. Or, during the assessment, the chosen method can be adjusted due to new insights, e.g. on data-availability. Interaction with stakeholders can be beneficial, also in this phase, i.e. involvement by providing the required (inventory) data and expert knowledge about the system dynamics under consideration (Hossain et al., 2017; Teah et al., 2016). Involvement of stakeholders in the previous phases might ease the gathering of data, because they are already involved. Furthermore, being involved in the assessment might increase the chance they support the results and take them into consideration during the decision-making process.

Methods do not only differ regarding the themes they cover, but also regarding the indicators that are used to quantify the metrics for these themes. That is, two methods can cover the same themes, within the same system boundaries, but still use different indicators to estimate their importance (Kok and Zijp, 2016). Harmonization of methods on the level of indicators for comparable applications is an important activity, allowing assessments results to be comparable with other studies, also when this was not the original goal of the assessment. A good example of such a harmonization is USEtox (Rosenbaum et al., 2008): an initiative of scientist to harmonize the different methods that were available to explore and compare expected (eco)toxicity impacts in the life cycle assessments of products. Also from the policy realm such harmonization is initiated, e.g. the Product and Organization Environmental Footprint initiative of the European Commission (EC, 2013). On the level of method selection however, there can be tensions, between the benefits of harmonization (ease of application and communication) and the benefits of being situation specific (precision). Harmonization of method selection, i.e. select one sustainability assessment method to be applied in all situations, leads to comparability of results, growing databases and mutual experience with the method. On the other hand, it might lead to unsustainable situations, missing case-specific important issues, or lack of stakeholder support in the decision-making process. A holistic method, applicable to all sustainability questions seems unachievable (Little et al., 2016; Sala et al., 2012b) and might also be unnecessary or even counter-productive. Basic principles can be universal: for example to consider the whole life cycle, or to consider a universal basic set of sustainability themes, but operationalization requires the delineation of situation specific system boundaries and data. Perhaps, the compromise would be to select a set of sustainability themes that should be considered in all sustainability assessments and harmonize the indicators and databases that are used for these themes. This basic set of themes is then always considered and case specifically complemented with extra issues. Such a general set should be chosen democratically, such as the sustainability development goals (UN, 2015a), but then more selective, in combination with techniques to narrow down the amount of indicators. The latter may consist of utilizing possible correlations between indicators (Steinmann et al., 2016), so that an assessment can focus on overarching, representative ones. Furthermore, the basic set of themes to be considered could be general, based on an analyses of drivers for global change (Rockström et al., 2009a; UN, 2009) or, it could be sector-specific (Broeren et al., 2017), or both. As such, harmonization and specification can be complimentary efforts to increase the impact of sustainability assessments for decision-making.

Part of the sustainability assessment is to present the results, of a wide range of different issues, in such a way that it can be understood by the stakeholders and in the further decision-making context, and acted on. In the presentation of results of any SA, it can a priori be expected, that there are three main possible outcomes of a sustainability assessment. Outcomes may highlight situations of:

- Multiple win-win: one scenario scores best on many or all aspects of the assessment; there is strong evidence to select this solution.
- Neutral: scenarios score comparably on most or all aspects of the assessments; there is no reason to prefer any measure to be taken, no improvement expected.
- Trade-offs: all scenarios have their benefits and their costs compared to the other scenarios: there

is reason to select a beneficial scenario, while trying to define counter-acting measures to reduce the trade-off potentials.

In case of trade-offs, an extra step might be taken so as to identify whether a solution can be adjusted in order to solve the trade-off. If not, the different benefits and costs of a planned solution scenario must be weighted in order to make a decision. Aggregation can be part of this weighting process. Aggregation leads to loss of information and comes with value judgements (Gasparatos and Scolobig, 2012). On the other hand, aggregation leads to presentation of the results that is easier to interpret (Özdemir et al., 2011). Hence, the aggregation method should be carefully and transparently selected and reported on, respecting the view of the stakeholders and should be added to the operationalization of the sustainability assessment identification key in the near future (Zijp et al., 2017).

Apart from evaluating results as discussed above, a specific struggle in the interpretation of the results is the use of environmental and social boundaries in the assessment. Most sustainability assessment methods do not use or specify the environmental boundaries that are needed for the interpretation of the results and are thus in fact comparative sustainable development assessment methods instead of sustainability assessment methods. On the other hand, the boundaries that are used in assessment methods are (partly) based on normative choices and thus subject to change. For example, the often quoted planetary boundaries are also chosen points on a dose-response relationship of increasing risk (Bass, 2009; Rockström et al., 2009a). Furthermore, an environmental boundary set at one spatial scale, e.g. the planetary, is not necessarily equally protective and thus also not equally applicable at lower spatial scales. Unsustainable situations at the local scale can occur without higher-scale boundary transgressions (Dearing et al., 2014; Molden, 2009). For example, the chemical footprint calculations in Chapter 5 show that at EU-scale emissions of organic substances do not exceed the environmental boundary (Zijp et al., 2014), but this does not mean that more local boundaries are also not transgressed. At the 'end of pipe' situation (local scale), there may be effects. Those may depend on local emissions, but also on local vulnerabilities. This is why in Chapter 6 the ecosystem vulnerability distribution is proposed to be used instead of a single value to characterize the boundary concept.

Various authors have proposed approaches to translate the planetary boundaries to a regional scale (Dearing et al., 2014; Teah et al., 2016). However, also within a region, dose-response relationships vary between locations (Brook et al., 2013), depending on both the location-specific environmental conditions and exposed biota (Zijp et al., in press). Hence, an environmental boundary derived for one situation might not be protective for a situation at a comparable spatial scale. This is not only true for environmental, but also for social and economic indicators (Raworth, 2012). One can wonder what the relevance of using boundaries in assessments is, when those boundaries vary across regions or in time. How do we know if we are in the 'safe and just space for humanity' when the boundaries are subject to change? Here, the concepts of absolute and relative assessment are relevant to consider again. For example, the different options used as boundary in the chemical footprint calculations in Chapter 5 result in different numeric expressions on the magnitudes of the footprints and, consequently, in different absolute interpretations (Zijp et al., 2014). The relative interpretation of the results did not change: the impacts of pesticide use in the river basin reduced over time, in line with the policy measures that were taken to that end. The 'relative' interpretation, and the conclusion that policy measures apparently helped improvement of ecosystem conditions is informative, without certainty on the precise position and definition of the environmental boundary. Thus, a relevant conclusion on sustainable development does not necessarily require a boundary. A boundary helps to identify which themes might need attention, it helps to delineate a 'distance from target'(Pope et al., 2004) and it is necessary mainly when the question of interest is: "Does the chosen solution reduces the impact enough for the situation to be called sustainable?".

Like many policy standards, 'flexible' boundaries are not meaningless. They can be used to reach a goal and provide direction. For example, our policies on the production and use of chemicals are geared to serve in the protection of ecosystem and human health against adverse effects of exposures to chemicals (EC, 2006). The standards used in these policies are based on laboratory experiments, fate models and safety margins: a scientific and a normative part. The societally important issues (here: human and ecosystem health protection) dictate the normative part. However, when another societal important issue gets entangled, for example recycling of resources, the normative part of the standard should not be treated as a static fact, but re-considered from the more multi-disciplinary point of view (Riding et al., 2015). The normative parts of boundaries and policy standards might be reconsidered when problems and solutions are analyzed from a multi- instead of a mono-disciplinary point of view (e.g. safety versus circularity) (Riding et al., 2015). Therefore, perhaps it is better not to present boundaries as fixed numbers (only), but also reveal the location-specific distributions found in the study area, e.g. the ecosystem vulnerability distribution shown in Chapter 6, and the consequences of normative choices. This is visualized in Fig. 7.2. The ecosystem vulnerability distribution can provide an indication of stressor level distributions that can be considered safe, not safe or at risk. However, this categorization reflects normative choices that were made in setting the boundaries. In the example of Fig. 7.2 these are i) the choice for a fraction of species richness lost when deriving the EVD and ii) the concrete place on the distribution of safe, at risk and not safe. The Figure shows that four alternative boundaries can be derived as a result of two normative choices in the procedure.

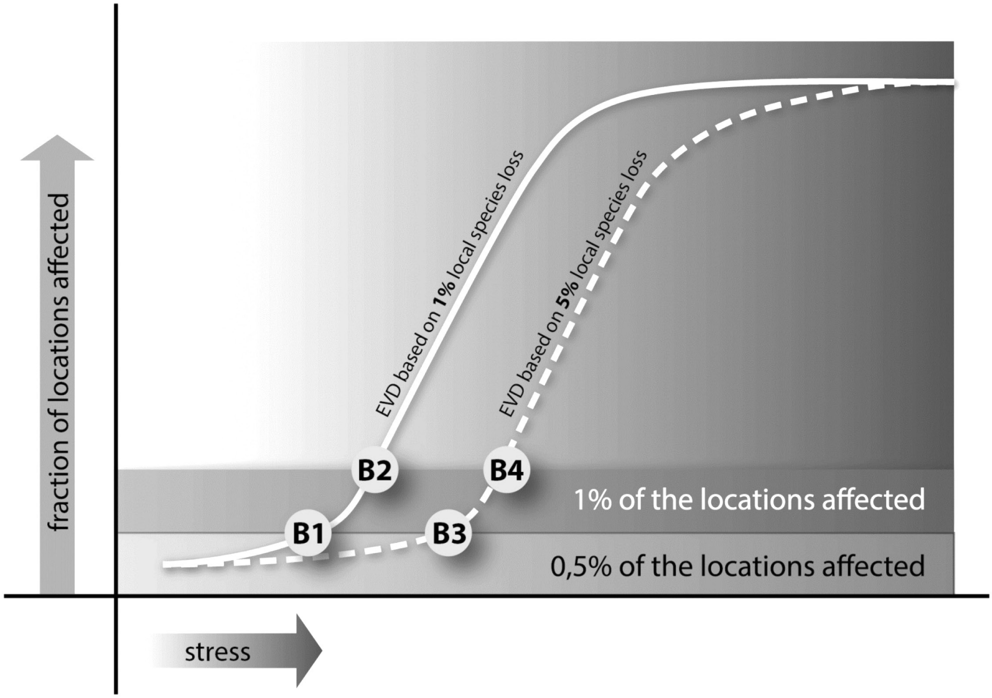


Fig. 7.2 The derivation and use of ecosystem vulnerability distributions requires value choices. These choices determine a boundary value, which may be positioned at B1 to B4, which differ as a consequence of (a) selecting an impact type (e.g. 1% or 5% species loss) and an impact level (e.g. 0.5% or 1% of the locations affected). The absolute outcomes of a sustainability assessment differ when choosing between B1, 2, 3 and 4.

7.6 Towards solution-focused sustainability assessments

Sustainability assessments are meant to support decisions that aim for sustainable development. Sustainable development is the effort to reach a social foundation for all within the carrying capacity of the earth. Reaching this goal is supported by (1) designing methods to support a shift of focus from the analysis of the problem towards an analysis of the problem and its optional solutions, (2) context specific and transparent sustainability assessment method selection, (3) knowledge about carrying capacities of ecosystems and (4) wisdom on how to operationalize this knowledge for decision-making. This thesis contributes to these challenges by designing, operationalizing and early-stage testing of 1) a procedure to execute solution-focused sustainability assessments; 2) an approach for SA method selection; and 3) two methods that enable stressor identification and ranking within a region with multiple human induced stressor types and magnitudes and multiple vulnerabilities of ecosystems, i.e. the chemical footprint method and the ecosystem vulnerability distribution. There are many ways by which these building blocks can be improved and deepened. The primary activities that follow from this thesis with regard to the further development and use the SfSA, the SAIK and the EVD are delineated as follows:

Application and evaluation of solution-focused sustainability assessment. Given the utility-improvement argument that triggered the solution-focused approach, and the wider need for practical and societally important multi-metric sustainability evaluations such as those described by Riding et al. (2014), there is a momentum to re-consider some environmental problems currently addressed as single-metric risk problems, to redefine them in the wider context of sustainability assessments, and to apply the solution-focused paradigm for finding sustainable solutions to those problems.

The most practical way to apply SfSA seems to include the approach and its underlying steps in the working procedures of organizations. That is for example in templates for project descriptions and in the project management courses. Then, after having gained some experience, the assumption of the increased utility of assessment outcomes for sustainability decisions should be evaluated. Do decision makers and stakeholders indeed value and use the sustainability assessment more, and is the quality and type of questions posed beneficially changing? And, do the approaches of solution-focused sustainability assessment and the sustainability assessment identification key lead to other assessments and other results that better represent the problem and its solutions? Can the outcomes better serve decision-making? In order to perform this evaluation, case studies should not only be performed, but also reported. As an example, the idea of Solution-focused Risk Assessment (SfRA) was discussed in literature (Davies, 2011; Finkel, 2011; Hope, 2011; Paustenbach, 2011; U.S. NAS, 2009), but so far case studies that refer to SfRA in literature are scarce. The current lack of a suite of reported case studies on the use of the solution-focused paradigm makes it impossible to draw final conclusions on the utility of the concept. However, experience in the case studies of this thesis hold the promise of assessment results that are aligned with and used in decision-making.

The evaluations should report on the lessons learned with regard to the assessment procedures, the sustainability metrics used and the role of the sustainability expert and/or researchers. That is, experts that perform sustainability assessments operate at the interface of research, government, representatives and industry (Arellano, 2015). This requires insight in different socio-ecological systems (Brondizio and Le Tourneau, 2016; Folke et al., 2016) and the ability to participate in these systems rather than studying them (Marshall et al., 2017; Spruijt et al., 2014).

Continue to develop the sustainability assessment identification key. Deriving an identification key is an iterative process (Nickerson et al., 2013), especially for a fast evolving discipline as sustainability science. Bearing that in mind, a first design of an identification key, as presented in this thesis, should be seen as a first essential step towards a complete and widely applicable and applied key. The proposed categorisation of the identification key should be continuously challenged in the context of striving for further improvement. Is this the optimal way to do it? There are other categorizations available in literature, e.g. Fang et al. (2017). Furthermore, elements such as data availability could be added, although with care: method selection should in first instance be based on what is required and not on what is available. Also, the options to consider aggregation across metrics of sustainable development should be added to the identification key.

Another important improvement is a further clarification on what exactly is a method. The method exploration in Chapter 4, that was used for the online tool, includes methods that cover an inventory (collecting data), an impact assessment and an aggregation phase. These three phases however, could also be seen as different methods, that are interchangeable. For example, life cycle impact assessment methods (LCIA) are developed for the impact assessment phase of LCAs, but can also be used in assessments that do not include different life cycle stages. Furthermore, methods are often unique combinations of comparable indicators. For example, most methods that include climate change as a theme, use life cycle assessment methods to quantify it, also when the other themes in the method are not life-cycle focused. A further involvement of the identification key could be to analyse which 'mother-methods' the different methods constitute of and to make case-specific combinations of agreed on 'mother-methods'.

Boundaries as instruments not as goals. In this era of big-data policy standards and boundaries can be more and more based on field data in addition to data from laboratory tests. The distribution of vulnerabilities within a region can be assessed based on field data. Now it is time to discuss the meaning of these distribution for setting boundaries and policy standards, or, to skip that part and use distributions directly in sustainability assessments and policy. The operationalization and use of the EVD as proposed in this thesis requires further exploration as well. For example, is it necessary to take biotic interaction into account when deriving the EVD? and, are EVDs region specific or can general applicable EVDs be derived?

Finally, setting boundaries involves normative choices, e.g. how much risk is allowed and which safety factors should be used. These choices are often based on one goal, such as protection of human health. However, in situations with conflicting goals that require a decision for sustainable development, the frequently-used normative choice might need to be reconsidered. Therefore, a next step in the derivation of boundaries would be to make the impact of normative choices on the boundaries visible, and use them as trading zone in decision procedures, comparable with sensitivity analyses in life cycle assessments based on different value choices in the used life cycle impact assessment method (De Schryver et al., 2011).

References

Abt, E., Rodricks, J.V., Levy, J.I., Zeise, L., Burke, T.A., 2010. Science and Decisions: Advancing Risk Assessment. Risk Anal 30, 1028-1036.

Adams, R., Jeanrenaud, S., Bessant, J., Denyer, D., Overy, P., 2015. Sustainability-oriented Innovation: A Systematic Review. Int J Manag Rev, 1-26.

Adriaanse, A., 1993. Environmental Policy Performance Indicators. The Hague, The Netherlands

AKWA. 2001. Base document for the ten-years scenario on sediment management: "Sediment in sight" (Basisdocument Tienjarenscenario Waterbodems - Bagger in Beeld). Utrecht, The Netherlands: Advies en kenniscentrum waterbodems

Aldenberg, T., Jaworska, J.S., Traas, T.P., 2002. Normal species sensitivity distributions and probabilistic ecological risk assessment. in: Posthuma L., Suter G.W., Traas T.P., eds. Species sensitivity distributions in ecotoxicology. Boca Raton, FL, United States of America: Lewis Publishers

Allen, F.H., Tainter, J.A., Hoekstra, T.W., 2003. Supply-Side Sustainability. New York: Columbia University Press

Arellano, A., 2015. The mechanisms and markers of research quality for Think Tanks. www.researchtoaction.org (last visit: 15-03-2017)

Arnot, J.A., Brown, T.N., Wania, F., Breivik, K., McLachlan, M.S., 2012. Prioritizing chemicals and data requirements for screening-level exposure and risk assessment. Environmental Health Perspectives 120, 1565-1570.

Bailey, K.D., 1994. Typologies and Taxonomies - An introduction to Classification Techniques. : SAGE, Thousand Oaks, California

Bartón, K., 2015. Multi-Model Inference. R-package,

Bass, S., 2009. Planetary boundaries: keep off the grass. Nat Clim Change, Commentary.

Beketov, M.A., Kefford, B.J., Schafer, R.B., Liess, M., 2013. Pesticides reduce regional biodiversity of stream invertebrates. Proc Natl Acad Sci USA 110, 11039–11043.

Benson, J.F., 2003. What is the alternative? Impact assessment tools and sustainable planning. Impact Assess Proj Apprais 21, 261-265.

Bernhardt, E.S., Rosi, E.J., O Gessner, M., 2017. Synthetic chemicals as agents of global change. Front Ecol Environ 15, 84-90.

Beroya-Eitner, M.A., 2016. Ecological vulnerability indicators. Ecol Indic 60, 329–334.

Binder, C.R., Feola, G., Steinberger, J.K., 2010. Considering the normative, systemic and procedural dimensions in indicator-based sustainability assessments in agriculture. Environ Impact Assess Rev 30, 71-81.

Blok, K., Huijbregts, M., Roes, L., Van Haaster, B., Patel, M., Hertwich, E., Wood, R., Hauschild, M., Sellke, P., Antunes, P., Hellweg, S., Ciroth, A., Harmelink, M., 2013. A novel methodology for the sustainability impact assessment of new technologies. Utrecht, the Netherlands

Bocken, N.M.P., de Pauw, I., Bakker, C., van der Grinten, B., 2016. Product design and business model strategies for a circular economy. J Ind Prod Engin 33, 308-320.

Boekhold, A.E., 2008. Ecological risk assessment in legislation on contaminated soil in The Netherlands. Sci Tot Environ 406, 518-522.

Böhringer, C., Jochem, P.E.P., 2007. Measuring the immeasurable - A survey of sustainability indices. Ecol Econ 63, 1-8.

Bond, A., Morrison-Saunders, A., Pope, J., 2012. Sustainability assessment: The state of the art. Impact Assess Proj Apprais 30, 53-62.

Bond, A.J., Morrison-Saunders, A., 2011. Re-evaluating Sustainability Assessment: Aligning the vision and the practice. Environ Impact Assess Rev 31, 1-7.

Brack, W., Altenburger, R., Schüürmann, G., Krauss, M., López Herráez, D., van Gils, J., Slobodnik, J., Munthe, J., Gawlik, B.M., van Wezel, A., Schriks, M., Hollender, J., Tollefsen, K.E., Mekenyan, O., Dimitrov, S., Bunke, D., Cousins, I., Posthuma, L., van den Brink, P.J., López de Alda, M., Barceló, D., Faust, M., Kortenkamp, A., Scrimshaw, M., Ignatova, S., Engelen, G., Massmann, G., Lemkine, G., Teodorovic, I., Walz, K.H., Dulio, V., Jonker, M.T.O., Jäger, F., Chipman, K., Falciani, F., Liska, I., Rooke, D., Zhang, X., Hollert, H., Vrana, B., Hilscherova, K., Kramer, K., Neumann, S., Hammerbacher, R., Backhaus, T., Mack, J., Segner, H., Escher, B., de Aragão Umbuzeiro, G., 2015. The SOLUTIONS project: Challenges and responses for present and future emerging pollutants in land and water resources management. Sci Tot Environ 503-504, 22-31.

Bradbury, J.A., 1989. The policy implications of differing concepts of risk. Science, Technol Hum Values 14, 380-399.

Brand, E., Otte, P.F., Lijzen, J.P.A., 2007. CSOIL2000. An exposure model for human risk assessment of soil contamination. A model description. Bilthoven: RIVM

Brewer, G.D., Stern, P.C., 2005. Decision making for the environment: social and behavioral science research priorities. Washington, DC.: National Academies Press

Broeren, M.L.M., Zijp, M.C., Waaijers-van der Loop, S.L., Heugens, E.H.W., Posthuma, L., Worrell, E., Shen, L., 2017. Environmental assessment of biobased chemicals in early-stage development - A review of methods and indicators. Biofuels, Bioproducts & Biorefining, doi:10.1002/bbb.1772.

Brondizio, E.S., Le Tourneau, F.M., 2016. Environmental governance for all. Science 352, 1272-1273.

Brook, B.W., Ellis, E.C., Perring, M.P., Mackay, A.W., Blomqvist, L., 2013. Does the terrestrial biosphere have planetary tipping points? Trends Ecol Evol 28, 396–401.

Browne, D., O'Regan, B., Moles, R., 2012. Comparison of energy flow accounting, energy flow metabolism ratio analysis and ecological footprinting as tools for measuring urban sustainability: A case-study of an Irish city-region. Ecol Econ 83, 97-107.

Bruggemann, R., Carlsen, L., 2012. Multi-criteria decision analyses. Viewing MCDA in terms of both process and aggregation methods: Some thoughts, motivated by the paper of Huang, Keisler and Linkov. Sci Total Environ, 293-295.

Brundtland, G.H., 1987. Report of the world commission on environment and development: Our common future. Addressed to the United Nations.

Bu, Q., Wang, D., Wang, Z., 2013. Review of screening systems for prioritizing chemical substances. Crit Rev Environ Sci Technol 43, 1011-1041.

Burton, J.G.A., 2002. Sediment quality criteria in use around the world. Limnology 3, 65-76.

Buser, A.M., MacLeod, M., Scheringer, M., Mackay, D., Bonnell, M., Russell, M.H., DePinto, J.V., Hungerbühler, K., 2012. Good modeling practice guidelines for applying multimedia models in chemical assessments. Integr Environ Assess Manage 8, 703-708.

Carlsson, A., Hjelm, O., Baas, L., Eklund, M., Krook, J., Lindahl, M., Sakao, T., 2015. Sustainability Jam Sessions for vision creation and problem solving. J Clean Prod, 29-35.

Carof, M., Colomb, B., Aveline, A., 2013. A guide for choosing the most appropriate method for multi-criteria assessment of agricultural systems according to decision-makers' expectations. Agricultural Systems 115, 51-62.

CBS, PBL, Wageningen UR. 2015. www.compendiumvoordeleefomgeving.nl. Last visit: december 2015)

CEFIC. 2012. Cefic guidance specific environmental release categories (SPERCs) chemical safety assessments, supply chain communication and downstream user compliance. By CEFIC.

Clark, W.C., Dickson, N.M., 2003. Sustainability science: the emerging research program. Proc Natl Acad Sci USA 100, 8059–8061.

Clift, R., Sim, S., King, H., Chenoweth, J.L., Christie, I., Clavreul, J., Mueller, C., Posthuma, L., Boulay, A., Chaplin-Kramer, R., Chatterton, J., DeClerck, F., Druckman, A., France, C., Franco, A., Gerten, D., Goedkoop, M., Hauschild, M.Z., Huijbregts, M.A.J., Koellner, T., Lambin, E.F.B., Lee, J., Mair, S., Marshall, S., McLachlan, M., Milà i Canals, L., Mitchell, C., Price, E., Rockström, J., Suckling, J., Murphy, R., 2017. The Challenges of Applying Planetary Boundaries as a Basis for Strategic Decision-making in Companies with Global Supply Chains. Sustainability 9, 279.

Coopers & Lybrand. 1997. Sediment removal within reach. Towards a structural solution for contaminated sediments in municipal waterways. The Hague, the Netherlands: Coopers & Lybrand Management Consultants.

Coppens, L.J.C., van Gils, J.A.G., ter Laak, T.L., Raterman, B.W., van Wezel, A.P., 2015. Towards spatially smart abatement of human pharmaceuticals in surface waters: Defining impact of sewage treatment plants on susceptible functions. Water Research 81, 356-365.

Čuček, L., Klemeš, J.J., Kravanja, Z., 2012. A review of footprint analysis tools for monitoring impacts on sustainability. J Cleaner Prod 34, 9-20.

Daly, H., 1990. Toward some operational principles of sustainable development. Ecol Econ 2, 1-6.

Daly, H., 1993. Sustainable growth: an impossibility theorem. in: Daly H., Townsend K., eds. Valuing the Earth: Economics, Ecology Ethics. Cambridge, MA: MIT Press.

Davies, J.C., 2011. Commentaries on finkel perspective article, comment on "solution-focused" risk assessment. Hum Ecol Risk Assess 17, 788–789.

De Fooij, H., 2015. Wastewater as a resource; Strategies to recover resources from Amsterdam's wastewater; MSc thesis in Civil Engineering and Management, University of Twente: Faculty of Engineering Technology.

De Laender, F., Morselli, M., Baveco, H., Van den Brink, P.J., Di Guardo, A., 2015. Theoretically exploring direct and indirect chemical effects across ecological and exposure scenarios using mechanistic fate and effects modelling. Environ Int 74, 181-190.

De Lange, H.J., Sala, S., Vighi, M., Faber, J.H., 2010. Ecological vulnerability in risk assessment — A review and perspectives. Sci Tot Environ 408, 3871-3879.

De Ridder, W., 2005. SustainabilityA-Test Inception report: progress to date and future tasks MNP rapport 555000001.

De Ridder, W., Turnpenny, J., Nilsson, M., Von Raggamby, A., 2007. A framework for tool selection and use in integrated assessment for sustainable development. J Environ Assess Policy and Manag 9, 423-441.

De Schryver, A.M., Van Zelm, R., Humbert, S., Pfister, S., McKone, T.E., Huijbregts, M.A.J., 2011. Value choices in life cycle impact assessment of stressors causing human health damage. J Ind Ecol 15, 796-815.

De Snoo, G.R., Vijver, M.G., 2012. Bestrijdingsmiddelen en waterkwaliteit. Leiden, the Netherlands.

De Valk, E., Hollander, A., Zijp, M.C., 2016. Environmental impact of the food consumption in the Netherlands. Bilthoven, the Netherlands: Research institute for Public Health and the Environment.

De Zwart, D., 2005. Ecological effects of pesticide use in The Netherlands: modeled and observed effects in the field ditch. Integr Environ Assess Manage 1, 123-134.

De Zwart, D., Dyer, S.D., Posthuma, L., Hawkins, C.P., 2006. Predictive models attribute effects on fish assemblages to toxicity and habitat alteration. Ecol Appl 16, 1295-1310.

De Zwart, D., Posthuma, L., 2005. Complex mixture toxicity for single and multiple species: Proposed methodologies. Environ Toxicol Chem 24, 2665-2676.

De Zwart, D., Warne, A., Forbes, V.E., Posthuma, L., Peijnenburg, W., Van de Meent, D., 2008. Matrix and media extrapolation. in: Solomon K.R., Brock T., De Zwart D., Dyer S.D., Posthuma L., Richards S., Sanderson H., Sibley P., Van den Brink P.J., eds. Extrapolation practice for ecotoxicological effect characterization of chemicals. Boca Raton, FL, USA: CRC-Press

Dearing, J.A., Wang, R., Zhang, K., Dyke, J.G., Haberl, H., Hossain, M.S., Langdon, P.G., Lenton, T.M., Raworth, K., Brown, S., Carstensen, J., Cole, M.J., Cornell, S.E., Dawson, T.P., Doncaster, C.P., Eigenbrod, F., Flörke, M., Jeffers, E., Mackay, A.W., Nykvist, B., Poppy, G.M., 2014. Safe and just operating spaces for regional social-ecological systems. Global Environmental Change 28, 227-238.

Den Hollander, H., Van Eijkeren, J.C.H., Van de Meent, D., 2004. Multimedia mass balance model for evaluating the fate of chemicals in the environment. Bilthoven, the Netherlands: National Institute of Public Health and the Environment.

Dietz, S., Neumayer, E., 2007. Weak and strong sustainability in the SEEA: Concepts and measurement. Ecol

Econ 61, 617-626.
Dijkers, M.P., Hart, T., Tsaousides, T., Whyte, J., Zanca, J.M., 2014. Treatment taxonomy for rehabilitation: past, present, and prospects. Arch Phys Med Rehab 95, S6-16.
Dodgson, J.S., Spackman, M., Pearman, A., Phillips, L.D., 2009. Multi-criteria analysis: a manual. London: Department for Communities and Local Government.
Dollar, D., Kraay, A., 2000. Growth is Good for the Poor. Washington, DC: World Bank.
Drescher, K., Boedeker, W., 1995. Assessment of the combined effects of substances: The relationship between concentration addition and independent action. Biometrics 51, 716-730.
Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L.J., Sullivan, C.A., 2006. Freshwater biodiversity: Importance, threats, status and conservation challenges. Biol Rev Cambridge Philos Soc 81, 163-182.
EC-JRC. 2010. International Reference Life Cycle Data System (ILCD) Handbook - General guide for Life Cycle Assessment - Detailed guidance. Luxembourg: European Commission - Joint Research Centre - Institute for Environment and Sustainability.
EC. 2000. Directive no 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Off J Eur Communities: Legis.
EC. 2006. Directive no 1907/2006 of the European Parliament and of the Council of 18 december 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) no 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Off J Eur Communities: Legis.
EC. 2013. Commission recommendation of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organisations. Off. J. Eur. Union.
EC. 2014a. European Innovation Partnership 'Agricultural productivity and sustainability'. Brussels: European Commission: http://ec.europa.eu/agriculture/eip/index_en.htm, last visit: 15-03-2017.
EC. 2014b. Horizon 2020 – Work programma 2014-2015; Annex G. Technological readiness levels (TRL). Brussels: European Commission.
EC. 2015. Towards an EU Research and Innovation policy agenda for Nature-Based Solutions & Re-Naturing Cities. Final Report of the Horizon 2020 Expert Group on 'Nature-Based Solutions and Re-Naturing Cities'. Brussels: European Commission.
EEA. 2011. Hazardous substances in Europe's fresh and marine waters. An overview. Copenhagen: European Environmental Agency.
Ellen MacArthur Foundation. 2013. Towards the Circular Economy Cowes, Isle of Wight: Ellen MacArthur Foundation.
Ellis, E.C., 2015. Ecology in an anthropogenic biosphere. Ecol Monog 85, 287-331.
European Commission. 2016a. Moving towards a circular economy. http://ec.europa.eu/environment/circular-economy/, last visit: 15-03-2017.
European Commission, 2016b. Request to the European Chemicals Agency for research on human health effects of granulated rubber. Ref Ares(2016)2535167 Brussels,.
Fang, K., Heijungs, R., De Snoo, G.R., 2014. Theoretical exploration for the combination of the ecological, energy, carbon, and water footprints: Overview of a footprint family. Ecol Indic 36, 508-518.
Fang, K., Song, S., Heijungs, R., De Groot, S., Dong, L., Song, J., Iswanto Wiloso, E., 2017. The footprint's fingerprint: on the classification of the footprint family. Curr Op Environ Sust 23, 54-62.
Fiala, N., 2008. Measuring sustainability: Why the ecological footprint is bad economics and bad environmental science. Ecol Econ 67, 519-525.
Finkel, A.M., 2011. Solution-focused risk assessment: a proposal for the fusion of environmental analysis and

action. Hum Ecol Risk Assess 17, 754–787.

Finnveden, G., Moberg, A., 2005. Environmental systems analysis tools - An overview. J Cleaner Prod 13, 1165-1173.

Florin, M.J., Van Ittersum, M.K., Van de Ven, G.W.J., 2012. Selecting the sharpest tools to explore the food-feed-fuel debate: Sustainability assessment of family farmers producing food, feed and fuel in Brazil. Ecol Indic 20, 108-120.

Folke, C., Biggs, R., Norström, A.V., Reyers, B., Rockström, J., 2016. Social-ecological resilience and biosphere-based sustainability science. Ecology and Society 21.

Freeman, E., 2012. Presence-Absence Model Evaluation. R-package.

Friesema, I.H.M., Tijsma, A.S.L., Wit, B., Van Pelt, W., 2014. Registry data of food-borne infections and food poisoning in the Netherlands in 2014. Bilthoven, the Netherlands: Research institute for Public Health and the Environment.

Gaasbeek, A., Meijer, E., 2013. Handbook on a novel methodology for the sustainability impact assessment of new technologies. Report prepared within the EC 7th Framework. www.prosuite.org (website currently offline): University of Utrecht.

Gadella, M., 2011. Evaluatie Besluit bodemkwaliteit; Evaluation of Decision soil quality. the Hague, the Netherlands: Agency NL - Bodem+.

Gandhi, N., Diamond, M.L., Van de Meent, D., Huijbregts, M.A.J., Peijnenburg, W.J.G.M., Guinée, J., 2010. New method for calculating comparative toxicity potential of cationic metals in freshwater: Application to Copper, Nickel, and Zinc. Environ Sci Technol 44, 5195-5201.

Gasparatos, A., El-Haram, M., Horner, M., 2008. A critical review of reductionist approaches for assessing the progress towards sustainability. Environ Impact Assess Rev 28, 286-311.

Gasparatos, A., Scolobig, A., 2012. Choosing the most appropriate sustainability assessment tool. Ecol Econ 80, 1-7.

Gavrilescu, M., 2014. Biomass potential for sustainable environment, biorefinery products and energy. in: Visa I., ed. Sustainable Energy in the Built Environment Springer International Publishing.

Geng, Y., Sarkis, J., Ulgiati, S., Zhang, P., 2013. Measuring China's circular economy. Science 339, 1526-1527.

George, J.S., Dennis, A.R., Nunamaker, J.F., 1992. An experimental investigation of facilitation in an EMS decision room. Group Decis Negot 1, 57-70.

Giubilato, E., Zabeo, A., Critto, A., Giove, S., Bierkens, J., Den Hond, E., Marcomini, A., 2014. A risk-based methodology for ranking environmental chemical stressors at the regional scale. Environ Int 65, 41-53.

Gleeson, T., Wada, Y., Bierkens, M.F.P., Van Beek, L.P.H., 2012. Water balance of global aquifers revealed by groundwater footprint. Nature 488, 197-200.

Goedkoop, M., Heijungs, R., Huijbregts, M.A.J., De Schryver, A.M., Struijs, J., Van Zelm, R., 2009. ReCiPe 2008. A life cycle impact assessment method which comprises harmonised category indicators at the midpoint and the endpoint level.

Gorman, M.E., Jenkins, L.D., Plowright, R.K., 2012. Human interactions and sustainability. in: Cabezas H., Diwekar U., eds. Sustainability: Mulit-disciplinary perspectives: Bentham Science Publishers.

Goussen, B., Price, O.R., Rendal, C., Ashauer, R., 2016. Integrated presentation of ecological risk from multiple stressors. Sci Rep 6.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: Offering more than simple habitat models. Ecol Lett 8, 993-1009.

Hacking, T., Guthrie, P., 2008. A framework for clarifying the meaning of triple bottom-line, integrated, and sustainability assessment. Environ Impact Assess Rev 28, 73-89.

Hage, M., Leroy, P.N., 2007. Guidance on stakeholder participation [Leidraad stakeholder participatie voor het Milieu en Natuurplanbureau. Praktijkwijzer; in Dutch]. Bilthoven, the Netherlands: PBL (Environmental Planning Agency).

Hajer, M., Nillson, M., Raworth, K., Bakker, P., Berkhout, F., De Boer, Y., Rockström, J., Ludwig, K., Kok, M., 2015. Beyond Cockpit-ism: Four Insights to Enhance the Transformative Potential of the Sustainable Development Goals. Sustainability 7, 1651-1660.

Harbers, J.V., Huijbregts, M.A.J., Posthuma, L., Van de Meent, D., 2006. Estimating the impact of high-production-volume chemicals on remote ecosystems by toxic pressure calculation. Environ Sci Technol 40, 1573-1580.

Harder, R., 2015. Fresh perspectives on the assessment of sewage sludge management. Gothenburg, Sweden: Chalmers University of Technology.

Harmsen, J., 2004. Landfarming of polycyclic aromatic hydrocarbons and mineral oil contaminated sediments. Wageningen: Wageningen University.

Harmsen, J., Rietra, R.P.J.J., Groenenberg, J.E., Lahr, J., Van den Toorn, A., Zweers, H.J., 2012. Verspreiden van bagger op het land in klei- en veengebieden. Wageningen, the Netherlands: Alterra.

Haughton, G., 1999. Environmental justice and the sustainable city. J Plan Educ Res 18, 233-243.

Heijungs, R., Goedkoop, M., Struijs, J., Effting, S., Sevenster, M., Huppes, G., 2003. Towards a life cycle impact assessment method which comprises category indicators at the midpoint and the endpoint level. Report of the first project phase: Design of the new method. The Hague: Ministry of VROM.

Henning-De Jong, I., Van Zelm, R., Huijbregts, M.A.J., De Zwart, D., Van Der Linden, A.M.A., Wintersen, A., Posthuma, L., Van de Meent, D., 2008. Ranking of agricultural pesticides in the Rhine-Meuse-Scheldt basin based on toxic pressure in marine ecosystems. Environ Toxicol Chem 27, 737-745.

Hickel, J., 2015. The problem with saving the world. in: Jacobin Magazine, ed. Brooklyn, U.S.: Jacobin,

Hierderer, R., 2012. EFSA spatial data version 1.1; Data properties and processing. Publications Office of the European Union.

Hinlopen, E., Nijkamp, P., 1990. Qualitative multiple criteria choice analysis. The Dominant Regime Method. Quality & Quantity. The Netherlands.

Hitchcock, K., Panko, J., Scott, P., 2012. Incorporating chemical footprint reporting into social responsibility reporting. Integr Environ Assess Manage 8, 386-388.

Hoekstra, A.Y., Chapagain, A.K., Aldaya, M.M., Mekonnen, M.M., 2009. Water Footprint Manual. State of the art 2009. Water Footprint Network.

Hoekstra, A.Y., Wiedmann, T.O., 2014. Humanity's unsustainable environmental footprint. Science 344, 4.

Hofstee, C., Leenaers, H., 2002. Actief Beheer van de Waterbodem (ABW) in landelijk perspectief; Active management of sediments in rural perspective. Utrecht, the Netherlands: TNO.

Hollander, A., De Jonge, R., Biesbroek, S., Hoekstra, J., M.C., Z., in prep. Exploring solutions for healthy, safe and sustainable n3-PUFA consumption in The Netherlands. Submitted.

Hope, B.K., 2011. A commentary on dr. Finkel's proposal for solution-focused risk assessment. Hum Ecol Risk Assess 17, 790–794.

Hopwood, B., Mellor, M., O'Brien, G., 2005. Sustainable Development: Mapping Different Approaches. Sust Dev 13, 38-52.

Hossain, M.S., Dearing, J.A., Eigenbrod, F., Johnson, F.A., 2017. Operationalizing safe operating space for regional social-ecological systems. Sci Tot Env 584-585, 673–682.

Huang, I.B., Keisler, J., Linkov, I., 2011. Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends. Sci Total Environ, 3578-3594.

Hyman, S.E., 2010. The diagnosis of mental disorders: The problem of reification. Annu Rev Clin Psycho 6, 155-179.

Ibáñez-Forés, V., Bovea, M.D., Azapagic, A., 2013. Assessing the sustainability of Best Available Techniques (BAT): Methodology and application in the ceramic tiles industry. J Cleaner Prod 51, 162-176.

ICLEI. 2014. Sustainable Cities. By the International Council for Local Environmental Initiatives (ICLEI), ed. www.sustainablecities.eu, last visited: 15-03-2017.

IPCC. 2001. Climate change 2001: impacts, adaptation, and vulnerability. Contribution of working group II to the third assessment report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge.

Ippolito, A., Sala, S., Faber, J.H., Vighi, M., 2010. Ecological vulnerability analysis: A river basin case study. Sci Tot Env 408, 3880-3890.

IRGC. 2008. An Introduction to the IRGC Risk Governance Framework. Geneva, Switzerland: International Risk Governance Council.

Isbell, F., Loreau, M., 2013. Human impacts on minimum subsets of species critical for maintaining ecosystem structure. Basic Appl Ecol 14, 623-629.

IUCN, UNEP, WWF. 1980. World Conservation Strategy: Living Resource Conservation for Sustainable Development. Gland, Switzerland: IUCN.

Jeswani, H.K., Azapagic, A., Schepelmann, P., Ritthoff, M., 2010. Options for broadening and deepening the LCA approaches. J Cleaner Prod 18, 120-127.

Joumard, R., 2011. Environmental sustainability assessments: Towards a new framework. Int J Sustainable Soc 3, 133-150.

Kapo, K.E., Holmes, C.M., Dyer, S.D., De Zwart, D., Posthuma, L., 2014. Developing a foundation for eco-epidemiological assessment of aquatic ecological status over large geographic regions utilizing existing data resources and models. Environ Toxicol Chem 33, 1665-1677.

Karlsson, M., Gilek, M., Udovyk, O., 2011. Governance of complex socio-environmental risks: The case of hazardous chemicals in the baltic sea. Ambio 40, 144-157.

Kates, R.W., Clark, W.C., Corell, R., Hall, J.M., Jaeger, C.C., Lowe, I., McCarthy, J.J., Schellnhuber, H.J., Bolin, B., Dickson, N.M., Faucheux, S., Gallopin, G.C., Grübler, A., Huntley, B., Jäger, J., Jodha, N.S., Kasperson, R.E., Mabogunje, A., Matson, P., Mooney, H., Moore III, B., O'Riordan, T., Svedin , U., 2001. Sustainability science. Science 292, 641-642.

Kendal, S., Creen, M., 2007. An introduction to Knowledge Engineering. London: Springer

Kibwika, P., Wals, A.E.J., 2008. Descending the Ivory Tower and re-making higher education in the era of (un) sustainability. Feature articles. Wageningen, the Netherlands: CTA.

Kingsford, R.T., Biggs, H.C., Pollard, S.R., 2011. Strategic Adaptive Management in freshwater protected areas and their rivers. Biol Conserv 144, 1194-1203.

Klewitz, J., Hansen, E.G., 2014. Sustainability-oriented innovation of SMEs: A systematic review. J Clean Prod 65, 57-75.

Kok, L., Zijp, M.C., 2016. Tools for Sustainable Public Procurement. Bilthoven, NL: Public Research Institute for Health and the Environment (RIVM).

Kramer, P.R.G., Huiting, A.M., Beurskens, J.E.M., Aldenberg, T., 1997. Verkenning bodemkwaliteit regionale wateren. Huidige en toekomstige gehalten van PAK in slootbodems. Bilthoven, The Netherlands: National Institute for Public Health and the Environment.

Kramer, P.R.G., Van Dijk, S., Beurskens, J.E.M., 1998. Verkenning bodemkwaliteit regionale wateren. Huidige en toekomstige gehalten van metalen in slootbodems. Bilthoven, The Netherlands: National Institute for Public Health and the Environment.

Kruijne R., Van der Linden A.M.A., Deneer J., Groenwold J.G., Wipfler E.L., 2011. Dutch environmental risk indicator for plant protection products. Alterra-rapport 2250. Wageningen: Alterra - Wageningen UR.

Landis, W.G., Chapman, P.M., 2011. Well past time to stop using NOELs and LOELs. Integr Environ Assess Manage 7, vi-viii.

Laniak, G.F., Olchin, G., Goodall, J., Voinov, A., Hill, M., Glynn, P., Whelan, G., Geller, G., Quinn, N., Blind, M., Peckham, S., Reaney, S., Gaber, N., Kennedy, R., Hughesm, A., 2013. Integrated environmental modeling: A vision and roadmap for the future. Environ Mod Softw 39, 3-23.

Leip, A., Weiss, F., Westhoek, J.P., 2013. The nitrogen footprint of food products in the European Union. J Agri

Science 152, 20-33.
Liang, X., Van Dijk, M.P., 2016. Evaluating the interests of different stakeholders in Beijing wastewater reuse systems for sustainable urban water management. Sustainability 8, 1098.
Lind, E.A., Kanfer, R., Earley, P.C., 1990. Voice, control, and procedural justice: Instrumental and noninstrumental concerns in fairness judgments. J Pers Soc Psychol 95, 952-959.
Lind, E.A., Tyler, T.R., 1988. The social psychology of procedural justice. New York: Plenum Press.
Linder, M., Williander, M., 2015. Circular Business Model Innovation: Inherent Uncertainties. Bus Strat Environ, DOI: 10.1002/bse.1906.
Linkov, I., Satterstrom, F.K., Kiker, G., Batchelor, C., Bridges, T., Ferguson, E., 2006. From comparative risk assessment to multi-criteria decision analysis and adaptive management: Recent developments and applications. Environ Int 32, 1072-1093.
Little, J.C., Hester, E.T., Carey, C.C., 2016. Assessing and enhancing environmental sustainability – a conceptual review. Environ Sci Technol 50, 6830-6845.
MacLeod, M., Scheringer, M., McKone, T.E., Hungerbuhler, K., 2010. The state of multimedia mass-balance modeling in environmental science and decision-making. Environ Sci Technol 44, 8360-8364.
Marshall, N., Adger, N., Attwood, S., Brown, K., Crissman, C., Cvitanovic, C., De Young, C., Gooch, M., James, C., Jessen, S., Johnson, D., Marshall, P., Park, S., Wachenfeld, D., Wrigley, D., 2017. Empirically derived guidance for social scientists to influence environmental policy. PLoS ONE 12.
Mata, T.M., Martins, A.A., Neto, B., Martins, M.L., Salcedo, R.L.R., Costa, C.A.V., 2012. Lca tool for sustainability evaluations in the pharmaceutical industry. Mech Chem Eng Trans 26, 261-266.
Mattila, T., 2013. Input-output analysis of the networks of production, consumption and environmental destruction in Finland. Doctoral Dissertations Aalto University, 120p.
Maxwell, S.L., Fuller, R.A., Brooks, T.M., Watson, J.E.M., 2016. The ravages of guns, nets and bulldozers. Nature 536, 143-145.
Mayer, A.L., 2008. Strengths and weaknesses of common sustainability indices for multidimensional systems. Environ Int 34, 277-291.
McNie, E.C., 2006. Reconciling the supply of scientific information with user demands: an analysis of the problem and review of the literature. Environ Sci & Pol 10, 17-38.
Ministerie van Infrastructuur en Milieu. 2016. Rijksbrede programma Circulaire Economie. the Hague, the Netherlands: Dutch ministery of Infrastructuur and the Environment.
Ministry V&W. 1998. Vierde Nota Waterhuishouding. Regeringsbeslissing. Ministerie van Verkeer en Waterstaat, Den Haag.
Ministry V&W, Ministry VROM, Ministry LNV, IPO. 1997. Actief Bodembeheer Rivierbed. Omgaan met verontreinigd sediment in de grote rivieren. Beleidsnotitie ministerie van V&W, ministerie van VROM, ministerie van LNV, IPO, Den Haag.
Ministry VROM. 1988. Premises for risk management (annex to the Dutch Environmental Policy Plan). Lower House, session 1988-1989, 21137, no. 5.
Ministry VROM. 1990. Notitie Milieukwaliteitsdoelstellingen bodem en water (MILBOWA). the Hague, the Netherlands: Ministry VROM.
Ministry VROM, 1993. Regeling vaststelling klasse-indeling onderhoudsspecie. Staatscourant.
Ministry VROM, 1997. Besluit van 8 december 1997, houdende vrijstellingen van het stortverbod buiten inrichtingen (Besluit vrijstellingen stortverbod buiten inrichtingen).
Ministry VROM, Ministry V&W, Ministry LNV, Unie van Waterschappen, Inter Provinciaal Overleg, Gemeenten, V.N., 2003. Newsletter Sediment and Soil. the Hague, the Netherlands.
Mitchell, J., Pabon, N., Collier, Z.A., Egeghy, P.P., Cohen-Hubal, E., Linkov, I., Vallero, D.A., 2013. A decision analytic approach to exposure-based chemical prioritization. PLoS ONE 8, 1-13.
Molden, D., 2009. Planetary boundaries: The devil is in the detail. Nat Clim Change.

Moreira, J.M.L., Cesaretti, M.A., Carajilescov, P., Maiorino, J.R., 2015. Sustainability deterioration of electricity generation in Brazil. Energy Policy 87, 334–346.

Mulder, C., Boit, A., Mori, S., Vonk, J., Dyer, S., Faggiano, L., Geisen, S., González, A., Kaspari, M., Lavorel, S., Marquet, P., Rossberg, A., Sterner, R., Voigt, W., Wall, D., 2012. Distributional (in)congruence of biodiversity–ecosystem functioning. Adv Ecol Res 46, 1-88.

Municipality Utrecht. 2015. Grondstof is afval, beleidsnota 2015-2018. Utrecht, the Netherlands: Municipality Utrecht.

Ness, B., Urbel-Piirsalu, E., Anderberg, S., Olsson, L., 2007. Categorising tools for sustainability assessment. Ecol Econ 60, 498-508.

Neumayer, E., 1999. Weak Versus Strong Sustainability: Exploring the Limits of Two Opposing Paradigms. Cheltenham, UK: Edward Elgar Publishing.

Nickerson, R.C., Varshney, U., Muntermann, J., 2013. A method for taxonomy development and its application in information systems. Eur J Inform Syst 22, 336-359.

Niemeijer, D., de Groot, R.S., 2008a. A conceptual framework for selecting environmental indicator sets. Ecol Ind 8, 14-25.

Niemeijer, D., De Groot, R.S., 2008b. Framing environmental indicators: Moving from causal chains to causal networks. Environ Develop Sust 10, 89-106.

Nieuwenhuis, R., Ellen, G.J., 2007. 'Leven met Bagger'. Samen met belanghebbenden komen tot duurzaam sedimentbeheer; 'Living with sediments'. Developing sustainable sediment mangement with all stakeholders. Tijdschrift Bodem 4, 180-182.

Nykvist, B., Persson, A., Moberg, F., Persson, L., Cornell, S., Rockström, J., 2013. National Environmental Performance on Planetary Boundaries. Stockholm, SE: Swedish Environmental Protection Agency.

Ocké, M.C., Toxopeus, I.B., Geurts, M., Mengelers, M.J.B., Temme, L., Hoeymans, N., 2017. What is on our plate? Bilthoven, the Netherlands: Research Institute for Public Health and the Environment.

OECD. 2002. Sustainable Development Strategies: A Resource Book. Paris: Organisation for Economic Co-Operation and Development.

OECD, 2003. OECD environmental indicators - development, measurement and use. OECD, Paris.

Ohio EPA. 2006. Methods for assessing habitat in flowing waters: using the Qualitative Habitat Evaluation Index (QHEI). Columbus, OH: Ohio Environmental Protection Agency; Division of Water Quality Monitoring and Assessment.

Ohio EPA. 2014. Ohio 2014 integrated water quality monitoring and assessment report. OH, USA: Ohio EPA; Division of Surface Water.

Olsen, S.I., Christensen, F.M., Hauschild, M., Pedersen, F., Larsen, H.F., Tørsløv, J., 2001. Life cycle impact assessment and risk assessment of chemicals - A methodological comparison. Environ Impact Assess Rev 21, 385-404.

Oomen, A.G.e., De Groot, G.M.e., 2016. Evaluation of health risks of playing sports on synthetic turf pitches with rubber granulate. Research institute for Public Health and the Environment.

Osté, L.A., Wintersen, A., Ten Kate, E., Posthuma, L., 2008. Nieuwe normen waterbodems. Normen voor verspreiden en toepassen op bodem onder oppervlaktewater. Lelystad: RIZA.

OVAM. 2005. Ontwerp uitvoeringsplan bagger en ruimingsspecie; Draft actionplan sediments and specie. Flanders: Openbare Vlaamse Afvalstoffenmaatschappij, OVAM.

Özdemir, E.D., Härdtlein, M., Jenssen, T., Zech, D., Eltrop, L., 2011. A confusion of tongues or the art of aggregating indicators—Reflections on four projective methodologies on sustainability measurement. Renew Sust Energ Rev 15, 2385-2396.

Panko, J., Hitchcock, K., 2011. Chemical footprint, ensuring product sustainability. Air and Waste Management Association, 12-15.

Papeş, M., Gaubert, P., 2007. Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. Divers Distrib 13, 890-902.

Parris, T.M., Kates, R.W., 2003. Characterizing and measuring sustainable development. Ann Rev Environ Resources 28, 559-586.

Paulson, P., 2015. Gates foundation rallies the troops to attack UN development-goals http://www.humanosphere.org/world-politics/2015/05/gates-foundation-rallies-the-troops-to-attack-un-development-goals/ last visit: 15-03-2017.

Paustenbach, D.J., 2011. Comments on dr. finkel's paper on solution focused risk assessment (sfra). Hum Ecol Risk Assess 17, 807-812.

Pearson, R.G., Raxworthy, C.J., Nakamura, M., Peterson, A.T., 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr 34, 102-117.

Peijnenburg, W., De Groot, A., Jager, T., Posthuma, L., 2005. Short-term ecological risks of depositing contaminated sediment on arable soil. Ecotoxicol Environ Saf 60, 1-14.

Pellissier, L., Rohr, R.P., Ndiribe, C., Pradervand, J., Salamin, N., Guisan, A., Wisz, M.S., 2013. Combining food web and species distribution models for improved community projections. Ecol Evol 3, 4572-4583.

Persson, L., Breitholtz, M., Cousins, I., de Wit, A., MacLeod, M., McLachlan, M., 2013. Confronting unknown planetary boundary threats from chemical pollution. Environ Sci Technol 47, 12619–12622.

Peters, I., 2009. Folksonomies: Indexing and Retrieval in Web 2.0.: De Gruyter, Saur: Berlin.

Pilière, A., Schipper, A.M., Breure, A.M., Posthuma, L., De Zwart, D., Dyer, S.D., Huijbregts, M., 2014. Unraveling the relationships between freshwater invertebrates assemblages and interacting environmental factors. Freshw Sci 33, 1148-1158.

Pintér, L., Hardi, P., Martinuzzi, A., Hall, J., 2012. Bellagio STAMP: Principles for sustainability assessment and measurement. Ecol Indic 17, 20-28.

Pooley, S.P., Mendelsohn, J.A., Milner-Gulland, E.J., 2013. Hunting cown the chimera of multiple disciplinarity in conservation science. Conserv Biol 28, 22-32.

Pope, J., Annandale, D., Morrison-Saunders, A., 2004. Conceptualising sustainability assessment. Environ Impact Assess Rev 24, 595-616.

Posthuma, L., Bjørn, A., Zijp, M.C., Birkveld, M., Diamond, M.L., Hauschild, M.Z., Huijbregts, M.A.J., Mulder, C., Van de Meent, D., 2014. Beyond safe operating space: Finding chemical footprint feasible Environ Sci Technol 48, 6057–6059.

Posthuma, L., De Zwart, D., 2006. Predicted effects of toxicant mixtures are confirmed by changes in fish species assemblages in Ohio, USA, rivers. Environ Toxicol Chem 25, 1094-1105.

Posthuma, L., de Zwart, D., 2012. Predicted mixture toxic pressure relates to observed fraction of benthic macrofauna species impacted by contaminant mixtures. Environ Toxicol Chem 31, 2175-2188.

Posthuma, L., De Zwart, D., 2014. Species Sensitivity Distributions. Encyclopedia of Toxicology, 3rd edition: Elsevier Inc., Academic Press.

Posthuma, L., De Zwart, D., Wintersen, A., Lijzen, J., Swartjes, F., Cuypers, C., Van Noort, P., Harmsen, J., Groenenberg, B.J., 2006a. Beslissen over bagger op bodem. Deel 1. Systeembenadering, model en praktijkvoorbeelden. Bilthoven, the Netherlands: National Institute for Public Health and the Environment.

Posthuma, L., Eijsackers, H.J.P., Koelmans, A.A., Vijver, M.G., 2008. Ecological effects of diffuse mixed pollution are site-specific and require higher-tier risk assessment to improve site management decisions: A discussion paper. Sci Tot Environ 406, 503-517.

Posthuma, L., Lijzen, J., Otte, P., Zwart, D., Wintersen, A., Osté, L., Beek, M., Harmsen, J., Groenenberg, B., 2006b. Beslissen over bagger op bodem. Deel 3. Modellering van risico's na verspreiding bagger.

Posthuma, L., Suter II, G.W., Traas, T.P., 2002. Species Sensitivity Distributions in Ecotoxicology. Boca Raton, FL, United States of America: Lewis Publishers.

Prato, S., La Valle, P., De luca, E., Lattanzi, L., Migliore, G., Morgana, J.G., Munari, C., Nicoletti, L., Izzo, G., Mistri, M., 2014. The "one-out, all-out" principle entails the risk of imposing unnecessary restoration costs: A study case in two Mediterranean coastal lakes. Marin Pollut Bullet 80, 30-40.

PRé, CML, RUN, RIVM. 2013. ReCiPe 2008—a life cycle impact assessment method which comprises harmonized category indicators at the midpoint and the endpoint level. First edition (revised). Report I: Characterisation. in: Goedkoop M., Huijbregts M.A.J., Heijungs R., De Schryver A., Struijs J., Van Zelm R., The Hague, The Netherlands.

Raworth, K., 2012. A Safe and Just Space for Humanity: Can We Live Within the Doughnut? in: Oxfam, ed. Oxfam discussion paper. Oxford, UK.

Reed, M.S., 2008. Stakeholder participation for environmental management: A literature review. Biol Conserv 141, 2417-2431.

Rees, W.E., 1992. Ecological footprints and appropriated carrying capacity: what urban economics leaves out. Environment & Urbanization 4, 121-130.

Renn, O., 2008. Risk governance: Coping with uncertainty in a complex world. London, UK and Sterling, USA: Earthscan.

Renn, O., Klinke, A., Van Asselt, M., 2011. Coping with complexity, uncertainty and ambiguity in risk governance: A synthesis. Ambio 40, 231-246.

Renton, A., 2009. Suffering the science; Climate change, people and poverty. By Oxfam, Oxford, UK.

Riding, M.J., Herbert, B.M.J., Ricketts, L., Dodd, I., Ostle, N., Semple, K.T., 2015. Harmonising conflicts between science, regulation, perception and environmental impact: The case of soil conditioners from bioenergy. Environ Int 75, 52-67.

Rijkswaterstaat. 2013. Waterbase. http://www.rijkswaterstaat.nl/water/waterdata_waterberichtgeving/watergegevens/ last visit: december 2014.

Rijnland. 2007. Subsidieverordening Baggerkosten Rijnland; Subsidy regulations sediment management costs Rijnland. Hoogheemraadschap van Rijnland, the Netherlands.

Rittel, H.W.J., Webber, M.M., 1973. Dilemmas in a general theory of planning. Policy Sciences 4, 155-169.

RIVM. The e-tox database. The Dutch Institute for Public Health and the Environment (RIVM).

Robèrt, K.H., Schmidt-Bleek, B., Aloisi De Larderel, J., Basile, G., Jansen, J.L., Kuehr, R., Price Thomas, P., Suzuki, M., Hawken, P., Wackernagel, M., 2002. Strategic sustainable development - Selection, design and synergies of applied tools. J Cleaner Prod 10, 197-214.

Roberts, N.C., 2000. Wicked problems and network approaches to resolution. Int Publ Manage Rev 1, 1-19.

Robinson, L.A., Levy, J.I., 2011. The revolving relationship between risk assessment and risk management. Risk Anal 31, 1334–1344.

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., De Wit, C.A., Hughes, T., Van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J.A., 2009a. A safe operating space for humanity. Nature 461, 472-475.

Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin Iii, F.S., Lambin, E., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., Van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J., 2009b. Planetary boundaries: Exploring the safe operating space for humanity. Ecol Soc 14, 1-33.

Rodricks, J.V., Levy, J.I., 2013. Science and Decisions: Advancing Toxicology to Advance Risk Assessment. Toxicol Sci 131, 1-8.

Rorije, E., Verbruggen, E.M.J., Hollander, A., Traas, T.P., Janssen, M.P.M., 2011. Identifying potential POP and PBT substances. Development of a new persistence/bioaccumulation-score. Bilthoven: National Institute for

Public Health and the Environment.
Rosenbaum, R.K., Bachmann, T.M., Gold, L.S., Huijbregts, M.A.J., Jolliet, O., Juraske, R., Koehler, A., Larsen, H.F., MacLeod, M., Margni, M., McKone, T.E., Payet, J., Schuhmacher, M., Van De Meent, D., Hauschild, M.Z., 2008. USEtox - The UNEP-SETAC toxicity model: Recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment. Int J Life Cycle Assess 13, 532-546.
Rouwette, E.A.J.A., Vennix, J.A.M., van Mullekom, T., 2002. Group model building effectiveness: a review of assessment studies. System Dynamics Review 18, 5-45.
Sala, S., Ciuffo, B., Nijkamp, P., 2013. A meta-framework for sustainability assessment. Research Memorandum 2013-16: Free University of Amsterdam.
Sala, S., Farioli, F., Zamagni, A., 2012a. Life cycle sustainability assessment in the context of sustainability science progress (part 2). Int J Life Cycle Assess 18, 1686-1697.
Sala, S., Farioli, F., Zamagni, A., 2012b. Progress in sustainability science: lessons learnt from current methodologies for sustainability assessment: Part 1. Int J Life Cycle Assess 18, 1653–1672.
Sala, S., Goralczyk, M., 2013. Chemical footprint: a methodological framework for bridging life cycle assessment and planetary boundaries for chemical pollution. Integr Environ Assess Manage 9, 623-632.
Sayer, J., Sunderland, T., Ghazoul, J., Pfund, J.L., Sheil, D., Meijaard, E., Venter, M., Boedhihartono, A.K., Day, M., Garcia, C., Van Oosten, C., Buck, L.E., 2013. Ten principles for a landscape approach to reconciling agriculture, conservation, and other competing land uses. Proc Nat Acad Sci 110, 8349-8356.
Schäfer, R.B., Caquet, T., Siimes, K., Mueller, R., Lagadic, L., Liess, M., 2007. Effects of pesticides on community structure and ecosystem functions in agricultural streams of three biogeographical regions in Europe. Sci Tot Env 382, 272-285.
Schäfer, R.B., Kühn, B., Malaj, E., Köning, A., Gergs, R., 2016. Contribution of organic toxicants to multiple stress in river ecosystems. Freshw Biol 61, 2116–2128
Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., Van Nes, E.H., Rietkerk, M., Sugihara, G., 2009. Early-warning signals for critical transitions. Nature 461, 53-59.
Schipper, A.M., Posthuma, L., De Zwart, D., Huijbregts, M.A.J., 2014. Deriving field-based species sensitivity distributions (f-SSDs) from stacked species distribution models (S-SDMs). Environ Sci Technol 48, 14464-14471.
Schneider, A.G., Townsend-Small, A., Rosso, D., 2015. Impact of direct greenhouse gas emissions on the carbon footprint of water reclamation processes employing nitrification-denitrification. Sci Tot Env 505, 1166-1173.
Sexton, K., Linder, S.H., 2014. Integrated Assessment of Risk and Sustainability in the Context of Regulatory Decision Making. Environ Sci Technol 48, 1409-1418.
Siebert, H., 1982. Nature as a life support system. Renewable resources and environmental disruption. Zeitschrift für Nationalökonomie 42, 133-142.
Simkin, S.M., Allen, E.B., Bowmana, W.D., Clark, C.M., Belnap, J., Brooks, M.L., Cade, B.S., Collins, S.L., Geiser, L.H., Gilliam, F.S., Jovan, S.E., Pardo, L.H., Schulz, B.K., Stevens, C.J., Suding, K.N., Throop, H.L., Waller, D.M., 2016. Conditional vulnerability of plant diversity to atmospheric nitrogen deposition across the United States. Proc Natl Acad Sci USA 113, 4086-4091.
Singh, R.K., Murty, H.R., Gupta, S.K., Dikshit, A.K., 2012. An overview of sustainability assessment methodologies. Ecol Indic 15, 281-299.
Smetanová, S., Bláha, L., Liess, M., Schäfer, R.B., Beketov, M.A., 2014. Do predictions from Species Sensitivity Distributions match with field data? . Environ pol 189, 126-133.
Solomon, K.R., Baker, D.B., Richards, R.P., Dixon, K.R., Klaine, S.J., La Point, T.W., Kendall, R.J., Weisskopf, C.P., Giddings, J.M., Giesy, J.P., Hall, L.W., Williams, W.M., 1996. Ecological risk assessment of atrazine in North American surface waters. Environ Toxicol Chem 15, 31-76.
Solomon, K.R., Brock, T., De Zwart, D., Dyer, S.D., Posthuma, L., Richards, S., Sanderson, H., Sibley, P., Van den

Brink, P.J. eds. Extrapolation practice for ecotoxicological effect characterization of chemicals. Boca Raton, FL, USA: CRC-Press;5063 2008.

Solomon, K.R., Giesy, J.P., La Point, T.W., Giddings, J.M., Richards, R.P., 2013. Ecological risk assessment of atrazine in North American surface waters. Environ Toxicol Chem 32, 10-11.

Spijker, J., Mol, G., Posthuma, L., 2011. Regional ecotoxicological hazards associated with anthropogenic enrichment of heavy metals. Environ Geochem Hlth 33, 409-426.

Spruijt, P., Knol, A.B., Vasileiadou, E., Devilee, J., Lebret, E., Petersen, A.C., 2014. Roles of scientists as policy advisors on complex issues: A literature review. Environl Sci Policy, 16-25.

Stahl, C., Cimorelli, A., 2013. A demonstration of the necessity and feasibility of using a clumsy decision analytic approach on wicked environmental problems. Integr Environ Assess Manag 9, 17-30.

Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., De Vries, W., De Wit, C.A., Folke, C., Gerten, D., Heinke, J., Mace, G.M., Persson, L.M., Ramanathan, V., Reyers, B., Sörlin, S., 2015. Planetary boundaries: Guiding human development on a changing planet. Science 347, 736-747.

Steinmann, Z.J.N., Schipper, A.M., Hauck, M., Huijbregts, M.A.J., 2016. How many environmental impact indicators are needed in the evaluation of product life cycles? Environ Sci Technol 50, 3913–3919.

Stockwell, D.R.B., Peterson, A.T., 2002. Effects of sample size on accuracy of species distribution models. Ecol Model 148, 1-13.

Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K., Norris, R.H., 2006. Setting expectations for the ecological conditions of streams: The concept of reference conditions. Ecol Appl 16, 1267-1276.

STOWA. 2016. Levenscyclusanalyse grondstoffenfabriek. Producten uit de RWZI. STOWA, RvO, Grondstoffenfabriek.

Strempel, S., Scheringer, M., Ng, C.A., Hungerbuhler, K., 2012. Screening for PBT chemicals among the "existing" and "new" chemicals of the EU. Environ Sci Technol 46, 5680–5687.

Strona, G., Lafferty, K.D., 2016. Environmental change makes robust ecological networks fragile. Nat Comm 7.

Struijs, J., Van der Grinten, E., Aldenberg, T., 2010. Toxic pressure in the Dutch delta measured with bioassays. Trends over the years 2000-2009. Bilthoven, the Netherlands: National Institute for Public Health and the Environment (RIVM).

Svarstad, H., Petersen, L.K., Rothman, D., Siepel, H., Wätzold, F., 2008. Discursive biases of the environmental research framework DPSIR. Land Use Policy 25, 116-125.

Swartjes, F.A., 1999. Risk-based assessment of soil and groundwater quality in the Netherlands: Standards and remediation urgency. Risk Anal 19, 1235-1249.

Swartjes, F.A., 2011. Dealing with contaminated sites. From theory towards practical application.: Springer Science+Business Media.

Swartjes, F.A., Rutgers, M., Lijzen, J.P.A., Janssen, P.J.C.M., Otte, P.F., Wintersen, A., Brand, E., Posthuma, L., 2012. State of the art of contaminated site management in The Netherlands: Policy framework and risk assessment tools. Science of the Total Environment 427-428, 1-10.

Teah, H.Y., Akiyama, T., Carlos, R.S., Rayo, O.V., Khew, Y.T.J., Zhao, S., Zheng, L., Onuki, M., 2016. Assessment of downscaling planetary boundaries to semi-arid ecosystems with a local perception: A case study in the middle reaches of Heihe River. Sustainability 8.

Thabrew, L., Wiek, A., Ries, R., 2009. Environmental decision making in multi-stakeholder contexts: applicability of life cycle thinking in development planning and implementation. J Cleaner Prod 17, 67-76.

The University of Hertfordshire. Pesticide Properties Database.

Thywissen, K., 2006. Components of risk: a comparative glossary. Studies of the University Research Counsel Education (SOURCE): Publication Series of the United Nations University-Institute for Environment and Human Security (UNU-EHS).

Traverso, M., Asdrubali, F., Francia, A., Finkbeiner, M., 2012. Towards life cycle sustainability assessment: An

implementation to photovoltaic modules. Int J Life Cycle Assess 17, 1068-1079.
Tyler, T.R., Lind, E.A., 1992. A relational model of authority in groups. Adv Exp Soc Psychol 15, 115-191.
U.S. NAS, 1983. Risk Assessment in the Federal Government: Managing the Process. The National Academies Press.
U.S. NAS, 2009. Science and Decisions: Advancing Risk Assessment. Washington: The National Academies Press
U.S. NAS, 2011. Sustainability and the U.S. EPA. The National Academies Press.
Udo de Haes, H.A., Sleeswijk, A.W., Heijungs, R., 2006. Similarities, differences and synergisms between HERA and LCA - An analysis at three levels. Hum Ecol Risk Assess 12, 431-449.
UN, 2009. Millennium Ecosystem Assessment: Current state and trends 2001-2005.
UN, 2012. Outcome document of RIO+20 conference. United Nations
UN, 2015a. Transforming our world: the 2030 Agenda for Sustainable Development. By the General Assembly of the United Nations, documentnumber: A/RES/70/1
UN, 2015b. Sustainable Development Solutions Network.
USEPA, 2000. Estimation Programs Interface (EPI) Suite Version 4.11. Washington: U.S. Enironmental Protection Agency
USEPA, 2009. The national study of chemical residues in lake fish tissue. Washington, DC, U.S. : Environmental Protection Agency, Office of Water
Van de Meent, D., Aldenberg, T., Canton, J.H., Van Gestel, C.A.M., Slooff, W., 1990. Streven naar waarden. Achtergrondstudie ten behoeve van de nota 'Milieukwaliteitsnormering water en bodem'. Bilthoven: National Institute for Public Health and the Environment-670101001
Van den Bergh, J.C., Grazi, F., 2013. Ecological footprint policy? Land use as an environmental indicator. J Ind Ecol. 18(1),10-18.
Van den Bos, K., Wilke, H.A.M., Lind, E.A., 1988. When do we need procedural fairness? The role of trust in authority. J Person Soc Psych 75, 1449-1458.
Van der Horn, S., Meijer, J., 2015. The landscape approach. The concept, its potential and policy options for integrated sustainable landscape management. the Hague, the Netherlands: PBL Netherlands Environmental Assessment Agency.
Van der Laan, H., Van Opijnen, J., Van Houwelingen, G., Sousi, M., Zijp, M.C., Kok, L., Van Heeswijk, T., 2016. Search for the optimal remineralisation technology. Report pilot research on remineralisation 2014-2016. Gouda: Oasen N.V.
Van Dijk, S., Kramer, P.R.G., Beurskens, J.E.M., 1998. Prognose van de metaalgehalten in de landbodem onder invloed van het verspreiden van baggerspecie. Bilthoven, The Netherlands: National Institute for Public Health and the Environment.
Van Dijk, S., Kramer, P.R.G., Beurskens, J.E.M., 1999. Indicatie van de (eco)toxicologische risico's van metaalgehalten in de landbodem onder invloed van het verspreiden van baggerspecie. Bilthoven, The Netherlands: National Institute for Public Health and the Environment.
Van Eerdt, M., Van Dam, J., Tiktak, A., Vonk, M., Wortelboer, R., Van Zeijts, H., 2012. Evaluatie van de nota duurzame gewasbescherming. Netherlands Environmental Assessment Agency (PBL).
Van Goethem, T.M.W.J., Azevedo, L.B., Van Zelm, R., Hayes, F., Ashmore, M.R., Huijbregts, M.A.J., 2013. Plant Species Sensitivity Distributions for ozone exposure. Environm Pollut 178, 1-6.
Van Nieuwenhuijzen, A.F., Visser, C., Odegard, I.Y.R., Bergsma, G.C., Uijterlinde, C.A., Erp-Taalman-Kip, v., 2016. Life cycle analysis of recovered resource products from used water. Singapore International Water Week. Singapore.
Van Noort, P., Cuypers, C., Wintersen, A., De Zwart, D., Peijnenburg, W.J.G.M., Posthuma, L., Harmsen, J., Groenenberg, B.J., 2006. Beslissen over bagger op bodem. Deel 2. Onderbouwing stofgedragmodellering en voorspelde landbodemconcentraties na verspreiding bagger op land. Bilthoven, the Netherlands: National Institute for Public Health and the Environment.

Van Passel, S., Meul, M., 2012. Multilevel and multi-user sustainability assessment of farming systems. Environ Impact Assess Rev 32, 170-180.

Van Rossum, C.T.M., Fransen, H.P., Verkaik-Kloostermans, J., Buurma-Rethans, E.J.M., Ocké, M.C., 2011. Dutch national food consumption survey. 2007-2010 Diet of children and adults aged 7 to 69 years. Bilthoven, the Netherlands: Institute for Public Health and the Environment.

Van Steenwijk, J.M., Mol, G.A.J., 1996. Evaluation regarding current and expected water and sediment quality and quality criteria. [In Dutch: Toetsing huidige en verwachte water(bodem)kwaliteit aan de grenswaarden]. in: RIZA, ed. Lelystad, the Netherlands: RIZA

Van Zelm, R., Huijbregts, M.A., Harbers, J.V., Wintersen, A., Struijs, J., Posthuma, L., Van de Meent, D., 2007. Uncertainty in msPAF-based ecotoxicological effect factors for freshwater ecosystems in Life Cycle Impact Assessment. Integr Environ Assess Manage 3, 203-210.

Vaquer-Sunyer, R., Duarte, C.M., 2008. Thresholds of hypoxia for marine biodiversity. Proc Natl Acad Sci USA 105, 15452–15457.

Velten, S., Leventon, J., Jager, N., Newig, J., 2015. What Is Sustainable Agriculture? A Systematic Review. Sustainability 7, 7833-7865.

Versteegh, J., Dik, H., 2012. The quality of drinking water in the Netherlands in 2011. Bilthoven: RIVM, National Institute for Public Health and the Environment.

Veul, M.F.X., Muijs, B., Eijsackers, H.J.P., 2000. Sediment spreading on land. Unanswered science questions. Wageningen, the Netherlands: Program Integrated Soil Reserach (PGBO).

Villa, F., McLeod, H., 2002. Environmental vulnerability indicators for environmental planning and decision-making: Guidelines and applications. Environ Manage 29, 335–348.

Vindimian, E., 2001. The role of ecotoxicology for monitoring ecosystem health. Freshw For 16, 91-97.

VROM. 2003. Beleidsbrief Bodem; Policy letter Soil. Den Haag: Ministerie VROM.

Waas, T., Hugé, J., Block, T., Wright, T., Benitez-Capistros, F., Verbruggen, A., 2014. Sustainability assessment and indicators: Tools in a decision-making strategy for sustainable development. Sustainability 6, 5512-5534.

Warmbaugh, J.F., Woodrow Setzer, R., Reif, D.M., Gangwal, S., Mitchell-Blackwood, J., Arnot, J.A., Jolliet, O., Frame, A., Rabinowitz, J., Knudsen, T.B., Judson, R.S., Egeghy, P., Vallero, D., Cohen Hubal, E.A., 2013. High-throughput models for exposure-based chemical prioritization in the ExpoCast project. Environ Sci Technol 47, 8479-8488.

Wezenbeek, J., Walthaus, H.H.J., Pruijn, M., Oste, L., De Boer, P.D., De Graaf, R.J., Lijzen , J.P.A., Posthuma, L., Wintersen, A., Van Eijsden, G.G., 2007. Ken uw (water)bodemkwaliteit, de risico's inzichtelijk. Bilthoven, the Netherlands: Grontmij.

Wheater, H., Evans, E., 2009. Land use, water management and future flood risk. Land Use Policy 26, Supplement 1, S251-S264.

Winnebeck, K.H., 2011. An abbreviated alternatives assessment process for product designers: a children's furniture manufacturing case study. J Clean Prod 19, 464-476.

Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dormann, C.F., Forchhammer, M.C., Grytnes, J.A., Guisan, A., Heikkinen, R.K., Høye, T.T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.C., Normand, S., Öckinger, E., Schmidt, N.M., Termansen, M., Timmermann, A., Wardle, D.A., Aastrup, P., Svenning, J.C., 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. Biol Rev 88, 15-30.

Wrisberg, N., Udo de Haes, H.A., Triebswetter, U., Eder, P., Clift, R., 2000. Analytical tools for environmental design and management in a systems perspective: Centre of Environmental Science, Leiden University.

Xiong, W., Li, J., Chen, Y., Shan, B., Wang, W., Zhan, A., 2016. Determinants of community structure of zooplankton in heavily polluted river ecosystems. Sci Rep 6, 11.

Yoe, C., 2011. Principles of risk analysis: decision making under uncertainty: CRC-Press.

Zijp, M.C., Heijungs, R., Van der Voet, E., Van De Meent, D., Huijbregts, M., Hollander, A., Posthuma, L., 2015.

An identification key for selecting methods for sustainability assessments. Sustainability 7, 2490-2512.
Zijp, M.C., Huijbregts, M., Schipper, A.M., Mulder, C., Posthuma, L., in press. Identification and ranking of environmental threats using ecosystem vulnerability distributions. Submitted.
Zijp, M.C., Posthuma, L., Devilee, J., Wintersen, A., Swartjes, F., 2016. Definition and use of Solution-focused Sustainability Assessment: a novel approach to generate, explore and decide on sustainable solutions for wicked problems. Environ Int 91, 319-331.
Zijp, M.C., Posthuma, L., Van De Meent, D., 2014. Definition and applications of a versatile chemical pollution footprint methodology. Environ Sci Technol 48, 10588–10597.
Zijp, M.C., Van der Laan, H., 2015. Life Cycle Assessment of two alternative drinking water production schemes at Oasen-Kamerik. Bilthoven: RIVM.
Zijp, M.C., Waaijers-van der Loop, S.L., Heijungs, R., Broeren, M.L.M., Peeters, R., Van Nieuwenhuijzen, A.F., Shen, L., Heugens, E.H.W., Posthuma, 2017. Method selection for sustainability assessments: the case of recovery of resources from waste water. J Environ Manag 197, 221-230.
Zoeteman, B.C.J., 2012. Sustainable development drivers. The role of leadership in government, business and NGO performance.: Edward Elgar Publishing.
Zoeteman, K., 2001. Sustainability of nations. Tracing stages of sustainable development of nations with integrated indicators. Int J Sust Dev World Ecol 8:2, 93-109.

Summary

The United Nations' sustainable development goals (SDGs) are meant to establish a social foundation for humanity within the carrying capacity of the earth. Sustainable development concerns all activities that support meeting that situation. Decisions that aim to reach the sustainable development goals in practice require consideration of multiple interacting environmental and social issues and the anticipation on trade-offs between conflicting goals. For example, when reduction of poverty comes at the cost of increased emissions of greenhouse gasses, or when a reduction in greenhouse gas emissions comes at the cost of depletion of rare minerals and increased emissions of toxic substances. Furthermore, different stakeholders tend to have different views on how sustainability can be achieved, which complicates decision-making and implementation of sustainable development measures.

Sustainability Assessment (SA) is scientific support for decision-making that aims to reach a sustainable situation. SAs are meant to support sustainable development by providing insights in the potential social and environmental effects of existing human activities and solutions for situations that are thought to be unstainable. In practice, however, SAs typically focus on the extent of problems, with little emphasis on possible solutions. Reasons for this mismatch are the design of well-known assessment schemes that primarily focus on the delineation of the problem, the limited attention for SA method selection, and the role of environmental boundaries that define when a situation can be called sustainable.

In view of the importance of sustainable development decisions for reaching the SDGs, and given the aforementioned limitations of current SAs, the main goal of this thesis was to conceptually design, operationalize and test a solution-focused sustainability assessment framework. The operationalization consisted of a procedural and a methodological part. The procedural part focusses on designing and testing a comprehensive approach for executing a sustainability assessment process with a solutions-focused approach and a transparent SA method selection. Method selection is highly relevant, because it is the concrete translation of sustainable development. Different stakeholders tend to have different views on what sustainable development should encompass. In order to perform a SA that fits the decision context, these views should be taken into account when selecting a method for the SA. In practice however, SA methods are often selected based on expert experiences and method availability and applicability.

Associated to the use of SA outputs in decision-making is the role of environmental boundaries, which lately receives much attention of both scientists and politicians, triggered by the concept of planetary boundaries. Although the recent formulation of this concept has been formulated on a planetary basis, many environmental boundaries are of a local nature, and they can vary from region to region. Moreover, the transgression of one environmental boundary may follow from a combination of human activities, as holds for e.g. the effects of chemical mixtures on ecosystem integrity. Therefore, the procedural suggestions provided in this thesis are complemented with a methodological part that focusses on the quantification of variation in stress from multiple human activities within a region, and the quantification of variation in vulnerability of ecosystems to that stress. Effects of chemical mixtures on biodiversity are used as an example of both. Combining the stress and vulnerability estimates can be used to evaluate if the human activities within a region take place with or without exceeding the carrying capacity of the ecosystems in that region.

In this thesis, the case is built, starting from the design of a solution-focused sustainability assessment framework, with specific attention for SA method selection (Chapters 2, 3 and 4). Thereupon, the design of methods for net stressor quantification and the comparison of the outcomes with additionally designed regional boundaries are elaborated (Chapters 5 and 6). In all chapters, the approaches that were designed were

also tested and illustrated with case studies (Chapters 2 – 6). The chapter contents are detailed below. Note that the emphasis of the case studies and the boundary chapters (5 and 6) in this thesis is on environmental issues.

In Chapter 2 a framework was designed to improve the utility of SA outcomes for decision-making in situations that require sustainable development and which show the characteristics of so-called "wicked" environmental problems. That is, the situation involves multiple social and environmental issues, there are multiple stakeholder perspectives, there are trade-offs between conflicting goals and there is, therefore, no single optimally sustainable solution that can easily be recognized by all the stakeholders that are involved. The design of the Solution-focused Sustainability Assessment (SfSA) framework was inspired by existing frameworks from the realms of risk governance, adaptive management, solution-focused risk assessment and sustainability assessments. It consists of a sequence of steps that can be followed iteratively:

1. The pre-assessment; here the problem, its context and the available knowledge are explored, including an exploration of the stakeholders identities and their perspectives on the situation.
2. Finding solution scenarios; here brainstorm techniques are used, together with stakeholders, to find possible solution scenarios for the problem. Furthermore, a selection is made from all proposed scenarios for which the sustainability is assessed.
3. Set the rules for the sustainability assessment; here the view of the stakeholders on sustainability is transparently translated in methods to assess the sustainability of the selected solution scenarios.
4. Sustainability assessment (SA); here the sustainability assessment is performed, to yield comparative insights in the selected SA-metrics for the current situation and the selected solution scenarios.
5. Evaluate the SA results and choose a solution; here the results of the SA are used in the decision-making procedure and one or more of the solutions are selected for implementation.
6. Evaluate if the situation has indeed optimized; here, based on monitoring and input of the stakeholders, the situation after implementing the solutions is evaluated, and if needed a new SfSA cycle is started.

The innovative parts of the framework are the exploration of alternative solution scenarios upfront in the sustainability assessment process (step 2) and the explicit focus on SA method selection (step 3). During the procedure, iterations across steps is possible, and in practice iteration will often be necessary. For example, a comparative quick-scan sustainability assessment (an early check on step 4) can bring focus to select a set of promising, solution scenarios (step 2) for which a thorough assessment is performed. Furthermore, although the choice for a solution to be implemented is central in step 5 of the procedure, important choices are made in every step, which is why participation of the problem-owners and the stakeholders are important throughout the procedure. Active involvement of the experts that perform the SA in defining the key societal questions and in finding realistic solutions to minimize risk and optimize sustainability (step 1-3) results in assessments (step 4) that are expected to be closer to (or better applicable for) management decision support.

The SfSA process was applied to a case study concerning the sustainable management of contaminated sediments that are continuously formed in ditches in rural, agricultural areas. This problem is wicked, as disposal of contaminated sediment on adjacent land is potentially hazardous to humans, ecosystems and agricultural products. Non-removal would, however, reduce drainage capacity, followed by increased risks of flooding, while contaminated sediment removal followed by offsite treatment would imply high budget costs and soil subsidence. Application of the SfSA steps served in solving this problem. During the iterative and interactive process of finding solutions and quantifying SA metrics for them, the views of experts as well as of non-professionals changed from a substances and risk-oriented perception of 'almost always suspect sediments, not to be spread on adjacent land' to systems-level assessment in the tradition of Paracelsus: not the compound, but expected quantitative risk estimates of the compounds in the whole-system setting were interpreted to discern cases where risks and impacts could occur in the sediment-soil system. This formed the basis for implementing a novel sediment management policy.

In Chapter 3, the focus was on developing a framework that helps to make a well-informed choice for a sustainability assessment method that fits the decision context, triggered in Step 3 of the SfSA approach. Often, a systematic problem analysis that guides the choice of the SA method is not performed. A new approach was proposed and developed to support SA method selection on the basis of question articulation: the Sustainability Assessment Identification Key. The identification key was designed to lead its users – which may be an expert, but also the stakeholder group established according to the SfSA steps – through all important choices needed for a comprehensive question articulation for the sustainability assessment at hand. Subsequently, sustainability assessment methods that fit the resulting specifications following from the question articulation are suggested by the identification key. The key consists of five domains with four or five questions that each require specification. Three of the five domains determine which method is suited: i.e. system boundaries (e.g. what is the spatial scale?), theme selection (e.g. what is to be sustained?) and aggregation (e.g. what level of aggregation is required?). The two other domains, method design (e.g. should stakeholders be involved) and organizational restrictions (e.g. which capacity is available for data selection?), determine how the selected method is used. The key was tested retrospectively on a set of thirty case studies from literature. This revealed that most case studies do not transparently describe why an applied method was selected, resulting in, amongst others, a mismatch between the case description and the SA. It is expected that using the identification key helps to circumvent such problems. It can help to improve: (i) transparency in the link between the case study problem definition and its context and the SA method selection; (ii) consistency between questions asked and answers provided; and (iii) internal consistency in methodological design.

Subsequently, in Chapter 4, the aforementioned sustainability assessment identification key was further operationalized with a protocol that facilitates method selection, using the identification key, together with stakeholders. The protocol guides the experts and/or the stakeholders in their exploration of i) the decision context, ii) the different views of stakeholders and iii) the translation into specifications for the SA. In addition, an online tool is presented for SA method selection. This tool identifies assessment methods that meet the specifications obtained with the protocol, and so far contains 30 existing sustainability assessment methods. The protocol and tool were applied in a case study regarding the recovery of resources from domestic waste water. An SA was required to support strategic choices on which alternatives for resource recovery to focus on. In several iterations with the problem-owners and sustainability experts, a combination of SA methods was selected and the sustainability assessment was performed. Based on the SA results, it was concluded that the use of recovered resources from waste water leads to a reduction in environmental impacts and resource depletion for all options under evaluation. Secondly, not one of the alternative options to recover resources from waste water that were evaluated scored best on all selected themes. Hence, choices will require an evaluation of trade-offs and weighing of the results (which themes can be identified as dominant, to forward sustainable development most). However, thirdly, the number of optional techniques for resource recovery that were considered could be reduced based on the results. The assessment results have been forwarded to the first phase of the decision procedure of the water managers, so that they can contribute to strategic choices for sustainable resource recovery from waste water in the Netherlands. The availability of a tool for SA method selection does not diminish the role of the sustainability expert. The case study showed that expert knowledge is required to propose a consistent set of sustainability themes based on all input of the stakeholders and, when necessary, propose a combination of complementary methods if there is not one exact match.

The above described illustrations of the two new procedural elements that aim to support proper and useful sustainability assessments, i.e. the solution-focused sustainability assessment framework and the sustainability assessment identification key, show how insights can be gained on the relative sustainability performance of alternative solutions and the current situation. They however do not (yet) answer the question of whether a sustainable situation will be reached, because they did not include environmental and social boundaries as quantification of what a sustainable situation is. The following chapters focus on exactly that: the confrontation of stress from multiple human activities within a region with the carrying capacity of the environment.

Chapter 5 presents a model that was designed to quantify the potential impact of multiple substances within a region as a chemical footprint. Setting and using environmental boundaries or natural thresholds as indication of the carrying capacity of the ecosystem was part of the method. A natural threshold is a scientifically derived point at the stress-response curve that marks a change in response to stress. An environmental boundary is a point on the stress-response that is set in order to analyze whether situations are safe or not. The latter includes normative choices, related to decision-making under uncertainty: the higher the uncertainty on the position of a true threshold, and the higher the potential impact of transgressing a true threshold, the larger the distance between threshold and selected boundary. This study illustrates the use of an environmental boundary defined by a classical, protective policy choice, and a natural threshold defined by food web impact analyses, to determine the chemical footprint. Firstly, a boundary commonly applied in protective chemicals policies was adopted, and adapted for use in the context of footprint assessment. This boundary was defined at the mixture exposure level at which 95% of the species is protected against the impacts of chemicals on vital characteristics, like growth and reproduction. Secondly, a scientific threshold was used, provided by ecological food web models. Food web models were used to define which fraction of primary deletions (direct local extinction of species caused by a pressure) in a food web triggers secondary deletions (local extinction of species caused by the extinction of other species). The level of pressure at which no secondary (indirect) deletions occur in any of the studied food webs was used as an example of a threshold. The chemical footprint was then determined as the amount of water needed to dilute the impact to a level at which either of the two boundaries is not transgressed. Subsequently, this water volume can be compared to the amount of water present in the region. Two case studies were executed to test and illustrate the method. The first case study illustrated that the production and use of a large number of organic substances in Europe stay within the currently set policy boundaries for chemical pollution, evaluated for the region as a whole. The second case study showed that the use of pesticides in Northwestern Europe exceeded the set boundaries, while showing a declining trend over time in a period in which European pesticides regulation aim at sustainable pesticide use, but also exhibit a stabilization of the predicted impact for the more recent years. Furthermore, the method could be used to rank the relative importance of individual substances in causing the predicted impacts, showing that 75% to 85% of the substances do not significantly contribute to the net impact and that the largest part of the predicted impact is determined by respectively 6% (38 of the 630 organic substances) and 3% (8 of the 274 pesticides) of the substances. This information can be used to evaluate whether a sustainable situation is reached and if not what type of solutions to focus on to improve the water quality in Europe.

Policy standards and boundaries are conceptualized often as single numbers, as was also done in the previous Chapter. For chemicals, the use of policy boundaries is often based on experiments with a small set of species under laboratory conditions, whilst for other stressors field-based impact assessments may be leading in the derivation of the boundary condition. In reality, however, ecological insights imply that the responses of ecosystems are location-specific, depending on the assemblage of species locally present and the local environmental conditions. That is, it is likely that a regional boundary for a stressor, such as toxic pressure, can be envisaged as a distribution rather than a single value. This is the subject matter of Chapter 6. To acknowledge that impacts of stressors are commonly shaped by both the stressor (type and magnitude) as well as by the location-specific vulnerability of the exposed biota, a method was designed to derive field-based ecosystem vulnerability distributions (EVD) that, together with the distributions of stressor levels in the region, can be used to identify and rank stressors in a region and analyze if the human activities in the region are likely to lead to unsustainable situations with regard to those stressors. The derivation of an EVD for a stressor starts with the selection of reference sites within a selected geographical region that are in near pristine condition. Then, region-wide biomonitoring data are used to derive species-specific pressure-response models (SDMs, Species Distribution Models), which are subsequently used to describe the probability of occurrence of species at the reference sites given their location-specific environmental conditions. Then, per reference site, the stacked probability of occurrence was calculated. This is a strong indicator for the species richness at a location. Subsequently, to derive the EVD for a stressor, the change of the species richness in relation to the stressor for each reference site is modelled in silico using the stressor-response models (other stressors kept constant). A maximum allowable level of species loss is chosen, and for every reference site the stressor level that would result in that impact level is quantified. The boundaries derived in this way appeared to differ among the reference sites, as a direct consequence associated to the differences in reference assemblage composition and environmental conditions. This phenomenon implies the presence of a distribution of vulnerabilities (Ecosystem Vulnerability Distribution or EVD) across local assemblages for that stressor in the region of interest. The EVD can be used for stressor identification and ranking by comparing the distribution of the measured stressor values of all the sampling sites in the region with the EVD, with overlap signaling that stress exposure transgresses the vulnerability, implying stressor effects. The application of the EVD in a case study on fresh water ecosystems in the state of Ohio (U.S.A.) showed that of the potential stressors under evaluation (conductivity, hardness, habitat quality, pH, toxic pressure total nitrogen and total phosphorous), total phosphorous and physical habitat quality are the most likely causes of local impairments of water bodies in Ohio.

Finally, Chapter 7 elaborates on the contribution of this thesis to solutions-focused sustainability assessment. Sustainability assessments are meant to support decisions that aim for sustainable development. Sustainable development is the effort to reach a social foundation for all within the carrying capacity of the earth. Reaching this goal is likely supported by (1) designing methods to support a shift of focus from the analysis of the problem towards an analysis of the problem and its optional solutions, (2) context-specific and transparent sustainability assessment method selection, (3) knowledge about carrying capacities of ecosystems, and (4) how to operationalize this knowledge for decision-making. This thesis contributes to these challenges by designing, operationalizing and testing 1) a procedure for solution-focused sustainability assessments; 2) an approach for SA method selection; and 3) two methods that enable stressor identification and ranking within a region with multiple human induced stressor types and magnitudes and multiple vulnerabilities of ecosystems. These methods enable identification and ranking of stressors in a region, and provide an indication of the sustainability of the human activities given the region specific carrying capacity of the ecosystem.

Samenvatting

De Verenigde Naties hebben doelen gesteld voor duurzame ontwikkeling: de 'Sustainable Development Goals'. Deze doelen beogen een wereld waarin voor iedereen in de sociale basis behoeften wordt voorzien binnen de grenzen die het milieu ons biedt. Duurzame ontwikkeling is gedefinieerd als de activiteiten die bijdragen aan het realiseren van die wereld. Duurzame ontwikkeling vraagt keuzes waarbij vaak sprake is van afwenteling tussen tegenstrijdige doelen ('trade-offs'). Zo kan een armoede reducerende maatregel ten koste gaan van de uitstoot van meer broeikasgassen, terwijl het verminderen van emissies van broeikasgassen ten koste kan gaan van uitputting van zeldzame metalen of leiden tot meer emissies van toxische stoffen in het milieu. Een complicerende factor is daarnaast dat verschillende belanghebbenden als het gaat om duurzame ontwikkeling vaak uiteenlopende visies hebben over wat belangrijk is en wat niet. Dit maakt een duurzame afweging niet eenvoudig.

Duurzaamheidsanalyses worden uitgevoerd ter ondersteuning van besluitvormingsprocessen die duurzame ontwikkeling tot doel hebben. Ze zijn bedoeld om inschattingen te geven van de impact op mens, milieu en samenleving door bestaande activiteiten en door mogelijke oplossingen voor niet-duurzame situaties. In de praktijk blijkt echter dat duurzaamheidsanalyses vaak gericht zijn op de omvang van problemen, en dat er minder aandacht is voor de ontwikkeling die nodig is om uit die problemen te komen. Deze mismatch wordt veroorzaakt door de huidige aanpak van bekende duurzaamheidsanalyses, een beperkte aandacht voor de methodeselectie die voorafgaat aan duurzaamheidsanalyses, en de rol van sociale- en milieugrenzen om te bepalen of een situatie duurzaam is of niet.

In het licht van de noodzaak om duurzame keuzes te maken, zodat de 'Sustainable Development Goals' kunnen worden gerealiseerd, én gegeven de bovengenoemde beperkingen van duurzaamheidsanalyse, is het doel van dit proefschrift het ontwikkelen, operationaliseren en testen van een raamwerk voor het uitvoeren van oplossingsgerichte duurzaamheidsanalyses (DA's). Het proefschrift reikt hiervoor zowel procedurele als methodologische innovaties aan. Het procedurele deel is gericht op het aanpakken van duurzaamheidsvraagstukken met een prominente aandacht voor het vinden van oplossingen voor het duurzaamheidsvraagstuk en transparante, context-afhankelijke methodeselectie. De methode voor het uitvoeren van een duurzaamheidsanalyse wordt in de praktijk vaak bepaald door de 'toevallig' aanwezige expertise. Hierdoor komt het regelmatig voor dat analyses niet aansluiten bij de context van het besluitvormingsproces, de visie van de belanghebbenden (de stakeholders) op duurzame ontwikkeling en de verwachtingen van de betrokkenen.

De laatste jaren krijgt het afleiden van milieugrenzen ten behoeve van DA's steeds meer aandacht van zowel wetenschappers als beleidsmakers, onder meer aangewakkerd door publicaties betreffende de zogenoemde 'Planetary Boundaries'. Hoewel deze discussie zich richt op milieugrenzen op wereldschaal, zijn veel milieugrenzen in werkelijkheid vaak locatie-specifiek. Daarnaast wordt het wel of niet overschrijden van deze grenzen bepaald door een scala aan verschillende activiteiten. Daarom heeft dit proefschrift, naast het procedurele deel, ook een methodologisch deel dat zich richt op het kwantificeren van milieudruk door diverse gelijktijdige activiteiten in een regio en het kwantificeren van variatie in gevoeligheid van ecosystemen voor die milieudruk binnen een regio. Met voorbeelden wordt aangetoond hoe de combinatie van stress- en gevoeligheidsinschattingen wordt gebruikt om te bepalen of de milieudruk door activiteiten van de mens passen binnen de draagkracht van het ecosysteem.

De voor dit proefschrift ontwikkelde procedure voor oplossingsgerichte DA's en de specificaties voor de methodeselectie zijn beschreven in hoofdstuk 2 t/m 4. De methodes voor het confronteren van milieudruk met de gevoeligheid voor milieudruk in een regio staan in hoofdstuk 5 en 6. In alle hoofdstukken zijn de ontwikkelde methodes en procedures uitgewerkt aan de hand van casestudies. De focus van de case studies en de discussies over het gebruik van grenzen om te bepalen wat duurzaam is en wat niet ligt op de milieukant van duurzaamheid en minder op de sociale kant. Hieronder staat een samenvatting per hoofdstuk.

Hoofdstuk 2 van dit proefschrift beschrijft en illustreert het ontwikkelde raamwerk voor een betere ondersteuning van DA bij besluitvorming die is gericht op duurzame ontwikkeling in situaties met karakteristieken van zogenoemde 'wicked problems'. Dit zijn problemen waarbij rekening moet worden gehouden met een mix van onderling afhankelijke sociale- en milieuaspecten, met stakeholders die verschillend denken over duurzame ontwikkeling en 'trade-offs' tussen tegenstrijdige doelen. Dit impliceert vervolgens een grote kans dat er geen optimale keuze mogelijk is waar alle belanghebbenden eenvoudig mee akkoord zullen gaan. Het raamwerk is gebaseerd op bestaande literatuur over risicobeheersing, iteratief management, oplossingsgerichte risico analyse en duurzaamheidsanalyses. Het bevat de volgende stappen die met de nodige onderlinge iteraties worden gevolgd:

- Voorbereiding; in deze stap wordt het probleem, haar context en de beschikbare kennis erover verkend. Dit behelst ook een verkenning van de stakeholders en hun visies op de situatie.
- Oplossingen bedenken; hier worden brainstormtechnieken gebruikt om samen met de stakeholders mogelijke oplossingen te verzamelen. Vervolgens wordt samen bepaald welke van alle mogelijke oplossingen verder zullen worden uitgewerkt.
- Randvoorwaarden voor de duurzaamheidsanalyse bepalen; in deze stap wordt de visie van de verschillende stakeholders besproken en vertaald in randvoorwaarden voor de methodeselectie. Op basis van die randvoorwaarden wordt een methode wordt geselecteerd waarmee de oplossingsrichtingen zullen worden geanalyseerd.
- Duurzaamheidsanalyse; hier wordt de bijdrage aan duurzame ontwikkeling van de huidige situatie en verschillende oplossingen geanalyseerd.
- Evaluatie van de analyseresultaten en kiezen van een of meerdere van de oplossingen ter implementatie; hier worden de analyseresultaten gebruikt in het besluitvormingsproces.
- Evaluatie of de situatie is verbeterd; hier wordt gemonitord wat er is veranderd na de implementatie van de oplossing(en) en er wordt verkend of er een nieuwe ronde nodig is voor verdere duurzame ontwikkeling.

Het raamwerk is innovatief vanwege het vroegtijdig verkennen van oplossingen in het proces, én de bewuste aandacht voor het selecteren van een methode die goed past bij de situatie. Het doorlopen van het raamwerk zal in de meeste gevallen iteraties tussen de verschillende stappen vragen. Bijvoorbeeld, een snelle duurzaamheidsanalyse (stap 4) kan nodig zijn voor de selectie van oplossingen (stap 2) waarvoor vervolgens een grondige analyse wordt uitgevoerd. Daarnaast moet worden opgemerkt dat, hoewel in stap 5 de besluitvorming centraal staat, er in elke stap van de procedure belangrijke besluiten worden genomen. Dit is waarom het zo belangrijk is dat de probleemeigenaar en de belanghebbenden (de stakeholders) betrokken zijn bij elke stap van de procedure. Actieve betrokkenheid van de duurzaamheidsexperts die de analyse uitvoeren bij het definiëren van het probleem en de context en het formuleren van mogelijke oplossingen (stap 1-3) resulteert in analyses (stap 4) die naar verwachting beter aansluiten bij het besluitvormingsproces (stap 5).

Het raamwerk is geïllustreerd met behulp van een casestudie in retrospectief, namelijk het omgaan met licht verontreinigde bagger in Nederland. Het verspreiden van deze bagger op de aangrenzend land is potentieel gevaarlijk voor mens en ecosysteem, maar niet baggeren is geen optie vanwege het dichtslibben van de waterwegen, en het afvoeren en behandelen van bagger gaat gepaard met zeer hoge kosten en bodemdaling. Het toepassen van de stappen in het raamwerk resulteerde in een oplossing voor dit probleem.

Het perspectief van de betrokkenen veranderde tijdens het iteratief en interactief doorlopen van het proces om te komen tot oplossingen en het kwantificeren van de risico's van de oplossingen. In plaats van bagger per definitie als probleem te beschouwen werd nieuw beleid gevormd op basis van kwantitatieve risicobeoordeling. Daarin werd zowel rekening gehouden met de stofeigenschappen van mogelijke aanwezige verontreiniging als met de kans dat mens en milieu hieraan verhoogd worden blootgesteld bij verspreiding van bagger op aangrenzend land.

In hoofdstuk 3 is de focus het selecteren van duurzaamheidsmethoden die passen bij de context van de besluitvorming. Vaak wordt er geen systematische probleem analyse uitgevoerd voordat er een methode wordt gekozen. In dit hoofdstuk wordt een nieuwe aanpak voorgesteld voor methodeselectie op basis van vraagarticulatie: een determineersleutel voor DA methoden (sustainability assessment identification key), analoog aan bijvoorbeeld de werkwijze van een Flora voor het determineren van planten. De 'identification key' leidt de gebruikers (bijvoorbeeld de duurzaamheidsexpert of groep stakeholders) langs belangrijke keuzes die nodig zijn voor de selectie van een DA methode. Vervolgens reikt de identification key methoden aan die passen bij de gemaakte keuzes. Op basis van literatuur onderzoek worden voor de eerste versie van de identification key vijf domeinen onderscheiden met ieder vier of vijf vragen die specificatie vragen. Drie van de domeinen zijn bepalend bij de vraag welke methode te selecteren. Dat zijn de domeinen systeemgrenzen (bijvoorbeeld: wat is de ruimtelijke schaal van de analyse), duurzaamheidsthema's (bijvoorbeeld: wat moet er worden beschermd?) en aggregatie (bijvoorbeeld: tot op welk niveau moet worden geaggregeerd?). De andere twee domeinen, methode design (bijvoorbeeld: is stakeholder betrokkenheid belangrijk?) en organisatorische beperkingen (bijvoorbeeld: hoeveel capaciteit is er beschikbaar voor het verzamelen van data?), bepalen hoe een methode wordt toegepast. De 'identification key' is retrospectief getest op dertig casestudies uit de literatuur. De meeste casestudies waren niet of slechts gedeeltelijk transparant over de methodeselectie, die daardoor regelmatig niet aansloot bij de vraagstelling. Het gebruik van de identification key leidt naar verwachting tot verbeterd inzicht in de koppeling tussen de vraag en de geselecteerde methode, meer consistentie tussen de vraag en de methode en meer consistentie bij het toepassen van de gekozen methode.

In hoofdstuk 4 is vervolgens de 'identification key' verder geoperationaliseerd. Er is een protocol ontwikkeld dat het gebruik van de 'identification key' in interactie met stakeholders faciliteert. Het helpt duurzaamheidsexperts en stakeholders bij hun verkenning van de context, de verschillende visies van de stakeholders, en de vertaling daarvan in specificaties voor de methodeselectie. Daarnaast is een online-tool ontwikkeld die helpt met het selecteren van duurzaamheidsmethoden op basis van de specificaties die verkregen zijn met het protocol. Deze tool bevat momenteel 30 duurzaamheidsmethoden. Zowel het protocol als de tool zijn toegepast in een casestudie over het winnen en toepassen van grondstoffen uit afvalwater. Door middel van verschillende iteraties met de probleemeigenaren en duurzaamheidsexperts is een combinatie van methoden geselecteerd en is een duurzaamheidsanalyse uitgevoerd. Op basis van de resultaten kon worden geconcludeerd dat het gebruiken van uit afvalwater gewonnen grondstoffen leidt tot minder milieudruk en minder uitputting van grondstoffen. Daarnaast bleek dat geen van de mogelijke oplossingen het beste scoorde op alle duurzaamheidsaspecten. Een strategische keuze voor een of meerdere oplossingen zal dus gepaard gaan met een evaluatie van de trade-offs en het wegen van de verschillende belangen. Wel kon op basis van de analyse het aantal mogelijke oplossingen worden gereduceerd. Een van de oplossingen bleek niet realiseerbaar in de huidige markt. De resultaten zijn ingebracht bij de besluitvormingsprocedure van de Waterschappen, zodat ze kunnen bijdragen aan strategische keuzes voor grondstoffenwinning uit afvalwater in Nederland. Het toepassen van de tool doet niets af aan de rol van de expert bij het selecteren van een methode. De case studie liet zien dat duurzaamheidsexperts nodig zijn om brede kennis over duurzame ontwikkeling in te brengen, om op basis van de stakeholder input een consistente set aan duurzaamheidsthema's samen te stellen en om een optimale combinatie van methoden voor te stellen.

Hoofdstuk 5 presenteert een model voor het bepalen van de impact van mengsels van stoffen die door menselijk handelen in een regio in het milieu terechtkomen. Deze impact wordt uitgedrukt als chemische voetafdruk. Het bepalen en gebruiken van milieugrenswaarden is onderdeel van de aanpak. Daarvoor worden twee typen milieugrenzen gebruikt. Ten eerste een wetenschappelijk bepaald stressniveau: een significante knik in de milieudruk-effect-relatie. Ten tweede een beleidsgrens die bedoeld is om te bepalen of een situatie veilig of duurzaam is of niet. Deze beleidsgrens is zowel op empirische kennis alsook op subjectieve keuzes (zoals veiligheidsmarges) gebaseerd. In deze studie wordt het gebruik van beide typen milieugrenzen geïllustreerd.

Ten eerste is een door beleid aangereikte normatieve milieugrens gebruikt. Deze grens is gebaseerd op een definitie van 'geen impact' als de situatie waarin meer dan 95% van de soorten is beschermd tegen mogelijke effecten (zoals aantasten van de groei en reproductie) van de emissies van schadelijke stoffen door menselijk handelen. Als wetenschappelijke milieugrens is gebruik gemaakt van kennis die wordt aangereikt door ecologische voedselwebmodellen. Dit zijn modellen die beschrijven bij welke fractie van verdwenen soorten (directe lokale eliminatie van soorten door milieudruk) een secundaire eliminatie plaatsvindt (een soort verdwijnt door het verdwijnen van andere soorten waar die soort afhankelijk van was). De fractie van soorten waarbij geen enkele secundaire eliminatie plaatsvond in de 18 voedselwebben uit een bekende voedselwebstudie is genomen als grens. De chemische voetafdruk is de hoeveelheid water die nodig is om de verontreiniging door activiteiten in een regio te verdunnen tot een impact die onder de gekozen milieugrens ligt.

De methode is toegepast op twee casestudies. De eerste case studie liet zien dat de productie en het gebruik van organische chemicaliën in Europa samen leiden tot een impact die lager is dan de vastgestelde milieugrenzen. Een tweede case studie liet zien dat het gebruik van bestrijdingsmiddelen in Noordwest Europa wel leidt tot overschrijding van de vastgestelde milieugrenzen. De impact is afgenomen na het invoeren van de Europese richtlijnen voor pesticidegebruik, en meer recent is deze afname gestagneerd. Daarnaast kon de methode worden gebruikt om de bijdrage van afzonderlijke chemicaliën aan de totale impact te analyseren. Hieruit volgde dat 75% tot 85% van de stoffen niet significant (statistisch) bijdragen aan de impact, en dat het grootste deel van de impact terug te leiden is tot 6% (38 van de 630 organische stoffen) en 3% (8 van de 274 pesticiden) van de stoffen. Deze informatie kan goed worden gebruikt voor het vinden van oplossingen voor toxische druk door activiteiten in Europa.

Een milieugrens is vaak geoperationaliseerd als één getal, zoals in het voorgaande hoofdstuk. Voor chemische stoffen zijn deze grenzen vaak gebaseerd op experimenten met een standaard-set aan soorten onder laboratorium condities, terwijl bij voor andere drukfactoren velddata zijn gebruikt voor het afleiden van milieugrenzen. Veel studies vanuit de ecologie laten echter zien dat de gevoeligheid van ecosystemen voor stress locatie-specifiek is. De gevoeligheid hangt af van welke soorten aanwezig zijn en de locatie-specifieke toestand van het milieu. De milieugrens voor een drukfactor in een regio kan dus eigenlijk worden gezien als een verdeling in plaats van een getal. Daarom is in hoofdstuk 6 een voorstel gedaan voor een methode waarmee velddata kan worden gebruikt om ecosysteem gevoeligheidsverdelingen (EVDs) af te leiden voor een regio. De eerste stap bij het afleiden van EVDs is het vinden van een selectie referentielocaties in de regio. Om recht te doen aan alle mogelijke situaties in een regio moeten deze referentielocaties zo min mogelijk negatief zijn aangetast door menselijk handelen, en moeten de geselecteerde locaties samen de variatie aan ecosysteemtypen in de regio omvatten. Vervolgens worden biomonitoring data uit de hele regio gebruikt om soort-specifieke oorzaak-effect modellen te ontwikkelen. Voor de soortensamenstelling op de verschillende referentielocaties worden de resultaten van deze modellen samengenomen, en wordt per drukfactor geanalyseerd hoe de soortensamenstelling reageert op veranderingen in stressniveau. Vervolgens wordt voor het afleiden van EVDs een maximaal toelaatbaar verlies aan soortenrijkdom gekozen. Per referentielocatie wordt het stressniveau waarop dit verlies wordt bereikt vastgesteld. De verdeling van deze stressniveaus per stressor tussen de verschillende locaties is de EVD. De EVD kan vervolgens worden gebruikt om te bepalen of, en zo ja in welke mate, de verschillende stressoren een probleem zijn in de regio door de EVD te confronteren met de verdeling van gemeten stressniveaus in de hele regio. Toepassing van de EVD op zoetwater in Ohio (VS) liet zien dat van de stressoren die waren meegenomen in de studie (conductie, hardheid, fysieke habitatkwaliteit, pH, toxische druk, stikstof concentraties en fosfaat concentraties) de concentraties van fosfaat en de habitatkwaliteit de grootste bedreiging vormen voor een gezonde kwaliteit van zoetwaterecosystemen in Ohio.

Ten slotte wordt in hoofdstuk 7 de link gelegd tussen bovenstaande innovaties en het streven naar duurzame ontwikkeling. Duurzaamheidsanalyses zijn bedoeld om besluitvorming richting duurzame ontwikkeling te ondersteunen. Duurzame ontwikkeling is het creëren van een wereld waarin voor iedereen wordt voldaan aan de sociale basisbehoeften binnen de grenzen die het milieu ons biedt. Hieraan bijdragen vraagt (1) een verruiming van de focus van de omvang van het probleem naar mogelijke oplossingen (2) een context-specifieke en transparante methodeselectie (3) kennis over de draagkracht van het ecosysteem en (4) wijsheid over hoe deze inzichten kunnen bijdragen aan besluitvorming. Dit proefschrift draagt bij aan deze uitdagingen door het ontwikkelen, operationaliseren en testen van 1) een procedure voor het oplossingsgericht analyseren van duurzaamheidsvraagstukken; 2) een aanpak voor transparante en contextafhankelijke methodeselectie; en 3) twee methoden waarmee verschillende typen gelijktijdig voorkomende door de mens veroorzaakte stressfactoren kunnen worden geconfronteerd met de variatie aan gevoeligheid van ecosystemen binnen een regio. Deze methoden bieden een inschatting van welke stressoren welke impact hebben, en hoe volhoudbaar die impact is gezien de regio-specifieke draagkracht van het ecosysteem.

About the author

Michiel Christiaan Zijp was born in 1980 in Zeist, the Netherlands. He studied Environmental Science at the Vrije Universiteit in Amsterdam, starting in 1999 and graduating in 2005. After getting married in that same year he got employed at the Laboratory for Ecological Risk Assessment at the Dutch Research Institute for Public Health and the Environment. He worked as researcher and policy-advisor on the implementation of the European Water Framework Directive and Groundwater Directive and was secretary of two Dutch technical working groups, regarding groundwater and soil quality assessment and management. In 2012 he started the PhD work that resulted in this thesis, working in the RIVM Centre for Sustainability, Environment and Health established in 2013. In that period he contributed to other projects as project leader and researcher, amongst others on the topics of sustainable procurement, sustainable food consumption and sustainable drinking water production.

List of publications

Identification and ranking of environmental threats using ecosystem vulnerability distributions. MC Zijp, MAJ Huijbregts, A Schipper, C Mulder, L Posthuma (Submitted)

Identifying change agents for Corporate Sustainability integration. J. van den Berg, MC Zijp, W. Vermeulen, S. Witjes (Submitted)

Exploring solutions for healthy, safe and sustainable n3-PUFA consumption in The Netherlands. A Hollander, R de Jong, S Biesbroek, J. Hoekstra, MC Zijp (Submitted)

Method selection for sustainability assessments: the case of recovery of resources from waste water. MC Zijp, SL Waaijers-van der Loop, R Heijungs, MLM Broeren, R Peeters, A van Nieuwenhuijzen, L Shen, EHW Heugens, L Posthuma. Journal of Environmental Management (2017), 197, 221-230

Environmental assessment of biobased chemicals in early-stage development - A review of methods and indicators. MLM Broeren, MC Zijp, SL Waaijers-van der Loop, EHW Heugens, L Posthuma, E Worrell, L Shen. Biofuels, Bioproducts & Biorefining (in press)

ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. MAJ Huijbregts, ZJN Steinmann, PMF Elshout, G Stam, F Verones, M Vieira, MC Zijp, A Hollander, R van Zelm, The International Journal of Life Cycle Assessment (2017), 22(3), 138-147

The environmental sustainability of the Dutch diet. Background report to 'What's on our plate? Safe, healthy and sustainable diets in the Netherlands'. A Hollander, EHM Temme, MC Zijp. RIVM report 2016-0198 (2016)

Sustainable Public Procurement supporting tools, L Kok, MC Zijp. (Dutch) Tools voor Maatschappelijk Verantwoord Inkopen, RIVM Report 2016-0204 (2016)

ReCiPe 2016: A harmonized life cycle impact assessment method at midpoint and endpoint level Report I: Characterization, MAJ Huijbregts, ZJN Steinmann, PMF Elshout, G Stam, F Verones, M Vieira, A Hollander, MC Zijp, R van Zelm. RIVM Rapport 2016-0104 (2016)

Definition and use of Solution-focused Sustainability Assessment: A novel approach to generate, explore and decide on sustainable solutions for wicked problems. MC Zijp, L Posthuma, A Wintersen, J Devilee, F A Swartjes. Environment International (2016) 91, p. 319-331

Search for the optimal remineralisation technology. Report pilot research on remineralisation 2014-2016, H van der Laan, J van Opijnen, G van Houwelingen, M Sousi, MC Zijp, L Kok, T van Heeswijk. Oasen N.V. report (2016)

Environmental impact of the food consumption in the Netherlands. E. de Valk, A. Hollander, MC Zijp. RIVM Report 2016-0074 (2016)

Soil degradation in life cycle assessment (LCA), A. Hollander, MC Zijp, H. van Wijnen. (Dutch) Bodemdegradatie in levenscyclusanalyse (LCA). Een ruimtelijke indicator voor fosforuitputting op wereldschaal, Tijdschrift Bodem (2016), 26(1)

A spatially explicit LCA-indicator for P-depletion in agricultural soils. A Hollander, MC Zijp, H van Wijnen. RIVM Report 2015-0198 (2016)

Life Cycle Assessment of two drinking water production schemes. MC Zijp, H van der Laan. RIVM Rapport 2015-0209 (2016)

What effect does production have on the availability of fresh water? A Hollander, M Huijbregts, MC Zijp, F Verones. H2O Water Matters 2016

An Identification Key for Selecting Methods for Sustainability Assessments. MC Zijp, R Heijungs, E van der Voet, D van de Meent, MAJ Huijbregts, A Hollander, L Posthuma. Sustainability (2015) 7, 2490-2512

Life cycle assessment of a new drinking water production scheme, MC Zijp, H van der Laan. (Dutch) Levenscyclusanalyse van een nieuw drinkwaterproductieproces, tijdschrift H2O; 30 september 2015.

Sustainability aspects and nutritional composition of fish: evaluation of wild and cultivated fish species consumed in the Netherlands. M Seves, L Temme, M Brosens, MC Zijp, J Hoekstra, A Hollander. Climatic

11

12

Change (2015) DOI 10.1007/s10584-015-1581-1

L Posthuma, A Bjørn, MC Zijp, M Birkveld, ML Diamond, MZ Hauschild, MAJ Huijbregts, C Mulder, D. van de Meent. Beyond safe operating space: Finding chemical footprint feasible. Environ Sci Technol. (2014) 48(11), 6057–9.

Definition and Applications of a Versatile Chemical Pollution Footprint Methodology. MC Zijp, L. Posthuma, D. van de Meent. Environmental Science and Technology (2014) 48, 10588-10597.

The chemical footprint and solutions for polluted water, MC Zijp, L. Posthuma, D. van de Meent. (Dutch) De chemische voetafdruk en oplossingen voor watervervuiling. Tijdschrift H20; 17 november 2014

Environmental profit of sustainable procurement. A quick-scan of the minimum criteria, MC Zijp, D de Zwart. (Dutch) Milieuwinst van Duurzaam Inkopen. Een quick-scan van de minimumeisen. RIVM report 250005001 (2013)

Update of the risk assessment for groundwater bodies. MC Zijp, ACM de Nijs, HFR Reijnders, W Verweij, S Wuijts. (Dutch) Bijwerken van de karakterisering van grondwaterlichamen. RIVM report 607402001 (2011)

Proposals for trend assessment in groundwater for the WFD. W Verweij, MC Zijp, LJM Boumans, HFR Reijnders. (Dutch) Voorstellen voor trendbepaling in grondwater voor de KRW. RIVM report 607402002 (2011)

Prospective study of the drinking water supply in the Netherlands. S Wuijts, CH Büscher, MC Zijp, W Verweij, CTA Moermond, AM de Roda Husman, BH Tangena, A Hooijboer. (Dutch) Toekomstverkenning drinkwatervoorziening in Nederland. RIVM Report 609716001 (2011)

Region-specific Groundwater management in practice. FA Swartjes, J Valstar, MC Zijp, P van Beelen, PF Otte. (Dutch) Gebiedsgericht grondwaterbeheer in de praktijk. Gebiedsafbakening, aanpak bronzone, procedure voor monitoring, (risicogebaseerde) toetsing grondwaterkwaliteit, kosten-batenanalyse RIVM-report 607050010 (2011)

Clarify article 7.3 of the WFD for groundwater bodies. MC Zijp, S Wuijts, HHJ Dik. (Dutch) Uitwerking artikel 7.3 KRW voor grondwaterlichamen. Drinkwaterfunctie bij karakterisering en toestandbeoordeling van grondwaterlichamen. RIVM report 60730012 (2010)

Dutch guidance for assessing the chemical status of groundwater bodies. MC Zijp, P van Beelen, LJM Boumans, R van Ek, ACM de Nijs, W Verweij, S Wuijts. (Dutch) Protocol voor de beoordeling van grondwaterlichamen. RIVM report 607300011 (2009)

Drinking water in the river basin management plans of member states in the Rhine and Meuse river basins. S Wuijts, MC Zijp, HFR Reijnders (Dutch) Drinkwater in stroomgebiedbeheerplannen Rijn- en Maasoeverstaten. RIVM report 734301034 (2009)

Conceptual models for the Water Framework Directive and the Groundwater Directive. J Spijker, R Lieste, MC Zijp, ACM de Nijs. (Dutch) Conceptuele modellen voor de Kaderrichtlijn Water en de Grondwaterrichtlijn. RIVM report 607300010 (2009)

Large scale groundwater contamination and the WFD. MC Zijp, W Verweij, CW Versluijs (Dutch) Grootschalige grondwaterverontreiniging en de KRW. RIVM report 607701001 (2008)

Exemptions in the Water Framework Directive and the Groundwater Directive. Three cases in the Netherlands. MC Zijp, HFMW van Rijswick, M Wienhoven, ACM de Nijs, BJ Pieters, W Verweij (Dutch) Uitzonderingsbepalingen in de Kaderrichtlijn Water en Grondwaterrichtlijn. Drie casussen die in Nederland spelen. RIVM report (2008)

Methods for applying extension under the WFD concerning groundwater case studies. MC Zijp, AM Durand, AMA van der Linden, HJ van Wijnen, HFMW van Rijswick (Dutch) Methodiek voor toepassing fasering en doelverlaging op grondwater. RIVM report 607300002 (2007)

Residues of pesticides in groundwater. An analyses for the the WFD. AMA van der Linden, HFR Reijnders, MC Zijp en AM Durand-Huiting (Dutch) Residuen van gewasbeschermingsmiddelen in het grondwater. Een analyse voor de KRW. RIVM report 607310001 (2007)

A tiered procedure to assess risk due to contaminant migration in groundwater. PF Otte, MC Zijp, K Kovar, JPA Lijzen, FA Swartjes, AJ Verschoor. RIVM report 711701056 (2007)

Acknowledgements

Dankwoord

Dankbaar ben ik voor alle mensen die geholpen hebben bij het tot stand komen van dit werk.

De mannen die de mogelijkheid creëerden: **Ton Breure**, bedankt voor alle moeite die je hebt gedaan voor het creëren van de mogelijkheid om te promoveren binnen mijn RIVM werkzaamheden. **Dik van de Meent**, bedankt voor het initiëren en indienen van het project Meetbaar Duurzaam waarbinnen ik mijn ei kwijt kon als promovendus.

Het inspirerende team dat mij begeleidde: **Leo Posthuma**, jij bent een geschenk uit de hemel. Betrokken, zowel inhoudelijk als persoonlijk; Inspirerend, zowel je denkraam als hoe je met anderen omgaat; en verbindend, zowel inhoudelijk tussen disciplines als tussen individuen en groepen. Ik hoop dat ik nog lang met je mag oplopen. **Mark Huijbregts**, bedankt voor je aandacht. Die was vrijwel altijd raak, zowel je commentaren als je lof (het Tv dominee compliment zal ik niet snel vergeten). **Reinout Heijungs**, bedankt voor je begeleiding. In Leiden maakte jij mij wegwijs in de expanderende wereld van duurzaamheidmethoden. Zonder die beginselen was ik verdwaald.

De leidinggevenden die mij de ruimte gaven die ik nodig had om dit onderzoek te doen: **Aart Sterkenburg, Ton Breure, Manon Zwart, Theo Traas, Birgit Loos, Leo Posthuma en Evelyn Heugens**. Ik hoop dat dit werk ook voor jullie goede vruchten oplevert.

De collega's: Alle co-auteurs waarmee ik de hoofdstukken heb mogen ontwerpen, maar ook mijn dagelijkse collega's op het RIVM en bij mijn jaartje CML en mijn sporadische bezoeken aan de RUN. Dank voor jullie inspiratie, jullie hulp en samenzijn! **Dick de Zwart**, **Dik van de Meent en Ton van der Linden**, jullie wil ik speciaal noemen, omdat ik jullie alle drie regelmatig gevraagd heb om hulp en dat jullie belangeloos aanhaakten tot de desbetreffende vraag was opgelost. Daarnaast bewonder ik de innovatieve manier waarop jullie data transformeren in waardevolle informatie en kennis.

De mannen van 'het oog wil ook wat': Bedankt voor jullie hulp bij de 'finishing touch' van dit proefschrift. **Rob Proost**, voor het opmaken en uitzoeken wat de meest duurzame manier van printen was, **Gaetan Dirven**, voor het brainstormen en ontwerpen van plaatjes die anderen waarschijnlijk wél snappen, en **papa** voor het supermooie schilderij dat ik mocht gebruiken als cover.

De paranimfen: **Anne Hollander**, wij zijn gaan samenwerken in de periode dat ik ging promoveren. Ik heb veel bewondering voor hoe nuchter en enthousiast je ingewikkelde projecten oppakt en afmaakt. Je hebt mij flink wat werk uit handen genomen waardoor ik meer tijd overhield voor de artikelen die uiteindelijk in dit boekje terecht zijn gekomen. Bijzonder hoe weinig woorden we nodig hebben om elkaar uit te leggen wat we bedoelen. Ik hoop dat we samen nog veel duurzaamheidsprojecten mogen aanpakken. **Ton Vonk Noordegraaf,** ik vind het een voorrecht zo'n vriendelijke, intelligente, energieke en grappige vriend te hebben. Je hebt me regelmatig flink bemoedigd. Bemoediging die aankwam omdat hij komt van iemand die heel succesvol met hetzelfde bijltje gehakt had. Anne en Ton, bedankt dat jullie mijn paranimfen willen zijn. Ik beschouw het als een eer.

Secure base: Tot slot een grote dank voor mijn 'secure base'. De veilige basis waarvan uit ik avonturen kan aangaan. Dat is mijn mooie, dappere en lieve vrouw, **Charissa**. Mijn geweldige kinderen, **Noah, Rhodé en Kaleb**, ik ben ongelofelijk trots op jullie. Mijn **ouders**, bedankt voor de (speel)ruimte die jullie hebben en geven. En natuurlijk mijn vrienden, familie en de Oosterkerk. Ik ben me er bewust ben dat ik ongelofelijk bof met jullie allemaal.

Michiel C. Zijp works at the Dutch Institute for Public Health and the Environment (RIVM) as researcher and project manager on projects with a multi-disciplinary character. As expert in the Centre for Sustainability, Environment and Health he explores and assists in solving complex problems across the science-policy interface with a focus on sustainable development. Example projects are sustainable procurement, sustainable food consumption and sustainable drinking water production. In (t)his Ph.D. thesis he proposes and illustrates solutions to improve the utility of sustainability assessments. He provides theory and case studies on context-specific method selection and use of boundaries for our natural environment, with solution-focused sustainability assessments as core of the work.

UITLEG VOOR DUMMIES

ISBN 978-94-92477-52-1
9 789492 477521